Wiley Finance Series

Hedge Fund modelling and analysis using MATLAB®

PAUL DARBYSHIRE
DAVID HAMPTON

WILEY

Hedge Fund Modelling and Analysis Using MATLAB®

For other titles in the Wiley Finance series
please see www.wiley.com/finance

Hedge Fund Modelling and Analysis Using MATLAB®

Paul Darbyshire
David Hampton

WILEY

This edition first published 2014

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloging-in-Publication Data is available

A catalogue record for this book is available from the British Library.

ISBN 978-1-119-96737-8 (hbk) ISBN 978-1-119-96768-2 (ebk)
ISBN 978-1-119-96767-5 (ebk) ISBN 978-1-118-90502-9 (ebk)

Cover images reproduced by permission of Shutterstock.com

Set in 10/12pt Times by Aptara Inc., New Delhi, India
Printed in Great Britain by TJ International Ltd, Padstow, Cornwall, UK

To Mum & Dad

Thank you for all your love and support, as always.
P.D.

To Marie-Christine, Juliette and Antoine,
with my unconditional love.
D.H.

Contents

Preface

This book is a practical introduction to modelling and analysing hedge funds based on the MATLAB® technical computing environment. MATLAB® is a high-level language and interactive environment for numerical computation, visualisation and programming. Using MATLAB®, you can analyse data, develop algorithms and create models and applications. The language, tools and built-in maths functions enable you to explore multiple approaches and reach a solution faster than with spreadsheets or traditional programming languages, such as C/C++ or Java. MATLAB® is the foundation for all products, including Simulink® and can be extended with add-on products for a whole range of applications, including statistics, computational finance and optimisation.

MATLAB® and Simulink® are registered trademarks of The MathWorks, Inc.

MATHWORKS CONTACT INFORMATION

For MATLAB® and Simulink product information, please contact:

The MathWorks, Inc.
3 Apple Hill Drive
Natick, MA, 01760-2098 USA
Tel: 508-647-7000
Fax: 508-647-7001
E-mail: info@mathworks.com
Web: mathworks.com

1. MATLAB® SOURCE CODE

This book assumes a working knowledge of MATLAB® and an ability to implement built-in functions and have a familiarity with developing basic MATLAB® applications.

All MATLAB® source code used throughout the book has been tried and tested with the following MATLAB® version and Operating System:

MATLAB® Version: 8.1.0.604 (R2013a)
Operating System: Microsoft Windows 8 Version 6.2
Java Version: Java 1.6.0_17-b04 with Sun Microsystems Inc.
All MATLAB® source code is displayed in the following format.
In addition, the following MATLAB® add-on products were included in the installation:

Financial Toolbox Version 5.2
Optimisation Toolbox Version 6.4
Statistics Toolbox Version 8.3

```
%File: barchart.m
%import XL data
[~,dates,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');
n = size(returns,1); %# data points

figure; %create figure
sDate = datenum(dates(1)); %set start date for x-axis
eDate = datenum(dates(end)); %set end date for x-axis
xData = linspace(sDate,eDate,n);

bar(xData, returns); %plot bar chart
xlabel('Date'); %add x label
ylabel('RoR (%)'); %add y label
box off;
ytick = get(gca,'YTick'); %format axes
set(gca,'YTickLabel', sprintf('%.2f|',ytick))
datetick('x','yyyy','keeplimits');
```

Sample MATLAB® source code

If `%< . . . . >` is used at the top of any MATLAB® source code it indicates that part of the code has been omitted for simplicity.

```
%File: optimisation.m
%< . . . . >

%target return, linear constraints and bounds
rstar = 9.0; %target return
Aeq = [R';ones(1,N)];
beq = [rstar;1];
lb = zeros(N,1);
ub = ones(N,1)*0.5;
```

Sample MATLAB® source code with code omitted

2. MATLAB® USER-DEFINED FUNCTIONS

As well as making use of numerous MATLAB® built-in functions, there are many user-defined MATLAB® functions (prefixed with the letter "f") that extend functionality where necessary. All MATLAB® user-defined functions are presented in the following format:

```
%File: fStd.m
%STANDARD DEVIATION
function m2 = fStd(x,flag,f)
%---------------------------------------------------------------------
%x:      returns
%flag:   0 = sample, 1 = population
%f:      reporting frequency
%m2:     standard deviation (sample or population)
%---------------------------------------------------------------------

     m2 = std(x,flag).*sqrt(f);
end
```

Sample MATLAB® user-defined function

Please note that we do not give any warranty for completeness, nor do we guarantee that the code is error free. Any damage or loss incurred in the application of the MATLAB® source code, functions and concepts discussed in the book are entirely the reader's responsibility.

If you notice any errors in the MATLAB® source code or you wish to submit a new method as a user-defined MATLAB® function, algorithm, model, or some improvement to a method illustrated in the book, you are more than welcome.

3. HYPOTHETICAL HEDGE FUND DATA

Throughout the book there is constant reference to many hedge fund return series and factors. The 10 hedge funds and 15 factors used are all *hypothetical* and have been simulated by the authors as a unique data set for demonstration purposes only. The techniques and models used in the book can therefore be tested on the hypothetical data before being applied to real-life situations by the reader. The hypothetical data are nonetheless close to what would be expected in reality. The 10 funds are a mixture of several major hedge fund strategies, i.e. Commodity Trading Adviser (CTA), Long-Short Equity (LS), Global Macro (GM) and Market Neutral (MN) strategies as described in the table below:

10 hypothetical hedge funds

Hedge Fund	Abbreviation
Commodity Trading Advisor	CTA1, CTA2, CTA3
Long-Short Equity	LS1, LS2, LS3
Global Macro	GM1, GM2
Market Neutral	MN1, MN2

The 15 factors are a mixture of both passive and active indices as described in the table below:

15 hypothetical hedge fund factors

#	Beta Factors	Abbreviation
1	Passive Global Stock Index	PSDX
2	S&P 500 Equity Index	S&P 500 Index
3	Passive Global Bond Index	PBond DX
4	Passive Long Global Commodity Index	PCom DX
5	Passive Long US$ Index	PUSD DX
6	Risk-Free Rate	Rf
	Industry Reference Alternative Beta Factors	
7	Commodity Trading Advisor Index	CTA Index
8	Long-Short Equity Index	LS Index
	Fama French Carhart Factors	
9	Value minus Growth	Val - Gr
10	Small Cap minus Large Cap	SC - LC
11	Momentum	Mom
	Active Alternative Beta Factors	
12	Active Global Stock Futures Index	ASDX
13	Active Global Bond Futures Index	ABDX
14	Active Global Commodity Futures Index	ACDX
15	Active Global Foreign Exchange Futures Index	AFDX

4. BOOK WEBSITE

The official website for the book is located at:

www.darbyshirehampton.com

The website provides free downloads to all of the hypothetical data and MATLAB® source code as well as many other useful resources.

The authors can be contacted on any matter relating to the book, or in a professional capacity, at the following email addresses:

Paul Darbyshire: pd@darbyshirehampton.com
David Hampton: dh@darbyshirehampton.com

1 The Hedge Fund Industry

The global credit crisis originated from a growing bubble in the US real estate market which eventually burst in 2008. This led to an overwhelming default of mortgages linked to subprime debt to which financial institutions reacted by tightening credit facilities, selling off bad debts at huge losses and pursuing fast foreclosures on delinquent mortgages. A liquidity crisis followed in the credit markets and banks became increasingly reluctant to lend to one another causing risk premiums on debt to soar and credit to become ever scarcer and more costly. The global financial markets went into meltdown as a continuing spiral of worsening liquidity ensued. When the credit markets froze, hedge fund managers were unable to get their hands on enough capital to meet investor redemption requirements. Not until the early part of 2009 did the industry start to experience a marked resurgence in activity realising strong capital inflows and growing investor confidence. Nevertheless, this positive growth has since been slowed as a result of the on-going European sovereign debt crisis affecting the global economy.

The aftermath of the financial crisis has clearly highlighted many of the shortcomings of the hedge fund industry and heightened the debate over the need for increased regulation and monitoring. Nevertheless, it has since been widely accepted that hedge funds played only a small part in the global financial collapse and suffered at the hands of a highly regulated banking system.

Chapter 1 introduces the concept of a hedge fund and a description of how they are structured and managed as well as a discussion of the current state of the global hedge fund industry in the light of past and more recent financial crises. Several key investment techniques that are used in managing hedge fund strategies are also discussed. Chapter 1 aims to build a basic working knowledge of hedge funds, and along with an overview of hedge fund data sources in Chapter 2, arm the reader with the information required in order to approach and understand the more quantitative and theoretical aspects of modelling and analysis developed in later chapters.

1.1 WHAT ARE HEDGE FUNDS?

Whilst working for *Fortune* magazine in 1949, Alfred Winslow Jones began researching an article on various fashions in stock market forecasting and soon realised that it was possible to neutralise *market risk*[1] by buying undervalued securities and *short selling*[2] overvalued ones. Such an investment scheme was the first to employ a *hedge* to eliminate the potential for losses by cancelling out adverse market moves, and the technique of *leverage*[3] to greatly improve profits. Jones generated an exceptional amount of wealth through his *hedge fund* over the 1950s and 1960s and continually outperformed traditional money managers. Jones refused to register the hedge fund with the Securities Act of 1933, the Investment Advisers Act of 1940, or the Investment Company Act of 1940, the main argument being that the fund was a *private* entity and none of the laws associated with the three Acts applied to this type of investment. It was essential that such funds were treated separately to other regulated markets since the use of specialised investment techniques, such as short selling and leverage, was not permitted under these Acts, neither was the ability to charge performance fees to investors.

So that the funds maintained their *private* status, Jones would never publicly advertise or market the funds but only sought investors through word of mouth, keeping everything as secretive as possible. It was not until 1966, through the publication of a news article about Jones' exceptional profit-making ability, that Wall Street and *High Net Worth*[4] (HNW) individuals finally caught on and within a couple of years there were over 200 active hedge funds in the market. However, many of these hedge funds began straying from the original *market neutral* strategy used by Jones and employed other seemingly more volatile strategies. The losses investors associated with highly volatile investments discouraged them from investing in hedge funds. Moreover, the onset of the turbulent financial markets experienced in the 1970s practically wiped out the hedge fund industry altogether. Despite improving market conditions in the 1980s, only a handful of hedge funds remained active over this period. Indeed, the lack of hedge funds around in the market during this time changed the regulators' views on enforcing stricter regulation on the industry altogether. Not until the 1990s did the hedge fund industry begin to rise to prominence again and attract renewed investor confidence.

Nowadays, hedge funds are still considered private investment schemes (or vehicles) with a collective pool of capital only open to a small range of institutional investors and/or wealthy individuals and having minimal regulation. They can be as diverse as the manager in control of the capital wants to be in

[1] *Market risk* (or *systematic risk*) is the risk that the value of an investment will decrease due to the impact of various market factors, for example changes in interest and foreign currency rates.

[2] See Section 1.4.1.

[3] *Leverage* is the use of a range of financial instruments or borrowed capital to increase the potential return of an investment (see Section 1.4.2).

[4] A *High Net Worth* (HNW) individual (or family) is generally assumed to have investable assets in excess of $1 million, excluding any primary residence.

terms of the investment strategies and the range of financial instruments which they employ, including stocks, bonds, currencies, futures, options and physical commodities. It is difficult to define what constitutes a hedge fund, to the extent that it is now often thought in professional circles that a hedge fund is simply one that incorporates any *absolute return*[5] strategy that invests in the financial markets and applies traditional as well as non-traditional investment techniques. Many consider hedge funds to be within the class of *alternative* investments along with private equity and real estate finance that seek a range of investment strategies employing a variety of sophisticated investment techniques beyond the longer established traditional ones, such as *mutual funds*.[6]

The majority of hedge funds are structured as limited partnerships with the manager acting in the capacity of general partner and investors as limited partners. The general partners are responsible for the operation of the fund, relevant debts and any other financial obligations. Limited partners have nothing to do with the day-to-day running of the business and are only liable with respect to the amount of their investment. There is generally a minimum investment required by *accredited investors*[7] of the order of $250,000–$500,000, although many of the more established funds can require minimums of up to $10 million. Managers will also usually have their own personal wealth invested in the fund, a circumstance intended to further increase their incentive to consistently generate above average returns for both the clients and themselves. In addition to the minimum investment required, hedge funds will also charge a fee structure related to both the management and performance of the fund. Such fees are not only used for administrative and on-going operating costs but also to reward employees and managers for providing positive returns to investors. A typical fee basis is the so-called *2 and 20* structure which consists of a 2% annual fee (levied monthly or quarterly) based on the amount of Assets under Management (AuM) and a 20% performance-based fee, i.e. an incentive-oriented fee. The performance-based fee, also known as carried interest, is a percentage of the annual profits and only awarded to the manager when they have provided positive returns to their clients. Some hedge funds also apply so-called *high water marks* to a particular amount of capital invested such that the manager can only receive performance fees, on that amount of money, when the value of the capital is more than the previous largest value. If the investment falls in value, the manager must bring the amount back to the previous largest amount

[5] *Absolute return* refers to the ability of an actively managed fund to generate positive returns regardless of market conditions.

[6] *Mutual funds* are similar in structure to hedge funds but are subject to much stricter regulation and limited to very specific investments and strategies.

[7] An *accredited investor* is one with a net worth of at least $1 million or who has made $200,000 each year for the past two years ($300,000 if married with a spouse) and has the capacity to make the same amount the following year.

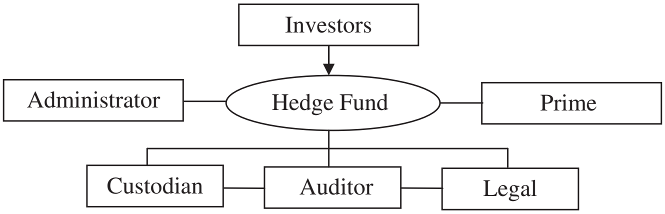


Figure 1.1 A schematic of the typical structure of a hedge fund

before they can receive performance fees again. A *hurdle rate* can also be included in the fee structure representing the minimum return on an investment a manager must achieve before performance fees are taken. The hurdle rate is usually tied to a market benchmark, such as LIBOR[8] or the one-year T-Bill rate plus a spread.

1.2 THE STRUCTURE OF A HEDGE FUND

In order for managers to be effective in the running of their business a number of internal and external parties covering a variety of operational roles are employed in the structure of a hedge fund as shown in Figure 1.1. As the industry matures and investors are requiring greater transparency and confidence in the hedge funds in which they invest, the focus on the effectiveness of these parties is growing, as are their relevant expertise and professionalism. Hedge funds are also realising that their infrastructure must keep pace with the rapidly changing industry. Whereas in the past, some funds paid little attention to their support and administrative activities, they are now aware that the effective operation of their fund ensures the fund does not encounter unnecessary and unexpected risks.

1.2.1 Fund Administrators

Hedge fund administrators provide many of the operational aspects of the successful running of a fund, such as compliance with legal and regulatory rulings, financial reporting, liaising with clients, provision of performance reports, risk controls and accounting procedures. Some of the larger established hedge funds use specialist in-house administrators whilst smaller funds may avoid this additional expense by outsourcing their administrative duties. Due to the increased requirement for tighter regulation and improved transparency

[8] LIBOR is the London Interbank Offered Rate, the average interbank interest rate at which a set of banks on the London interbank market are willing to lend to one another.

Table 1.1 Top five global administrators as of 2012

Administrator
State Street Alternative Investment Solutions
Citco Fund Services
BNY Mellon Alternative Investment Services
SS&C GlobeOp
Citi Hedge Fund Services

in the industry, many investors will only invest with managers that can prove a strong relationship with a reputable third-party administrator and that the proper processes and procedures are in place (see Table 1.1).

Hedge funds with offshore operations often use external administrators in offshore locations, to provide expert tax, legal and regulatory advice for those jurisdictions. Indeed, it is a requirement in some offshore locations (e.g. the Cayman Islands) that hedge fund accounts must be regularly audited. In these cases, administrators with knowledge of the appropriate requirements in those jurisdictions would fulfil this requirement.

1.2.2 Prime Brokers

The prime broker is an external party who provides extensive services and resources to a hedge fund, including brokerage services, securities lending, debt financing, clearing and settlement and risk management. Some prime brokers will even offer incubator services, office space and *seed* investment for start-up hedge funds. The fees earned by prime brokers can be quite considerable and include trade commissions, loan interest and various administration charges. Due to the nature of the relationship between the prime broker and a hedge fund, in particular being the counterparty to trades and positions, only the largest financial institutions are able to act in this capacity, for example Goldman Sachs, J P Morgan and Deutsche Bank. For this reason the prime brokerage market is relatively small and each prime broker tends to service a large number of hedge funds and therefore takes on an extremely high degree of risk. Some major restructuring occurred amongst prime brokers in 2008 and 2009, for example the acquisitions of Bear Stearns by J P Morgan and Merrill Lynch by Bank of America, and the takeover of Lehman Brothers by Barclays Capital. This resulted in a shift in market share from some former investment banks to commercial banks and saw the prime brokerage industry begin to consolidate. In order to alleviate investor concerns since the collapse of some major financial institutions, many fund managers are cautious in employing a single prime broker and prefer to subscribe to multiple prime brokers.

1.2.3 Custodian, Auditors and Legal

Hedge fund assets are usually held with a custodian, including the cash in the fund as well as the actual securities.[9] The custodian is normally a bank who will offer services, such as safekeeping the hedge fund assets, arranging settlement of any sales or purchases of securities and managing cash transactions.

The general structure of a hedge fund precludes them from the requirement to have their financial statements audited by a third party. However, in order to satisfy investors, many hedge funds have their accounts and financial reviews audited annually by an external audit firm. It is important that the auditing firm is seen to be independent of the hedge fund to give credence to their reports and services.

The legal structure of a hedge fund is designed to provide investors with limited liability, i.e. if a fund suffers a severe loss the maximum amount an investor can lose is the level of capital invested in the fund. They cannot be made liable for losses over this amount or any other outstanding debt or financial obligation. In addition, the legal structure is also chosen to optimise the tax status and legal liability of the hedge fund itself. To facilitate this, there are a small number of standard hedge fund structures, e.g. the master-feeder structure, which is adopted by a large number of funds. These comply with the legal requirements of the various jurisdictions where the hedge funds operate and obtain the optimal tax treatment. The master-feeder structure is a two-tier structure where investors invest through a feeder vehicle which itself invests in the hedge fund. There can be a number of feeder vehicles, located and domiciled in a number of different jurisdictions. Each can have a different legal form and framework. Depending on their tax status, investors can decide which feeder vehicle they wish to invest. As a general rule the tax regime of an investor will depend on the location of the investor, i.e. *on*- or *offshore*.[10]

1.3 THE GLOBAL HEDGE FUND INDUSTRY

After exceptional growth since 1998, total assets managed by the hedge fund industry peaked at $1.97 trillion in 2007. After the credit crunch and financial crisis of 2008, with well-publicised frauds and scandals as well as the collapse of several major financial institutions, the hedge fund industry suffered severe losses and investor loyalty. Not until the early part of 2009 did the industry start to experience a marked resurgence in activity realising strong capital inflows and growing investor confidence as shown in Figure 1.2. Strong rallies in global

[9] This is true except when the assets are used as *collateral* for gaining leverage. In these cases, the assets used as collateral are held by the prime broker. As most hedge funds use some degree of leverage, it is common for assets to be held by both custodians and prime brokers.

[10] *Onshore* (or *domestic*) locations include the US and UK, and to a lesser degree Switzerland and some other European countries. *Offshore* locations include the Cayman Islands, Bermuda, Bahamas, Luxembourg and Ireland.

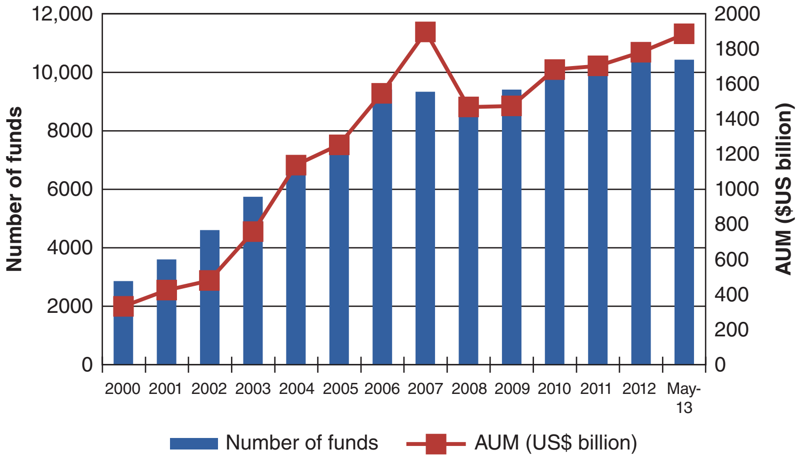


Figure 1.2 Growth in the global hedge funds industry since 2000
Source: Eurekahedge

markets in the last eight months of 2009 and the subsequent positive asset flows in 2010 aided the industry's recovery but growing concerns about the state of the global economy and the European sovereign debt crisis slowed down this recovery in 2011 and for most of 2012. It is estimated that the total amount of AuM in the global hedge fund industry at the end of May 2012 stood at $1.88 trillion.

Of the global hedge fund market, North American funds still remain the prominent market making up around two thirds of the global industry, followed, quite a way behind, by European and then Asian sectors (see Figure 1.3).

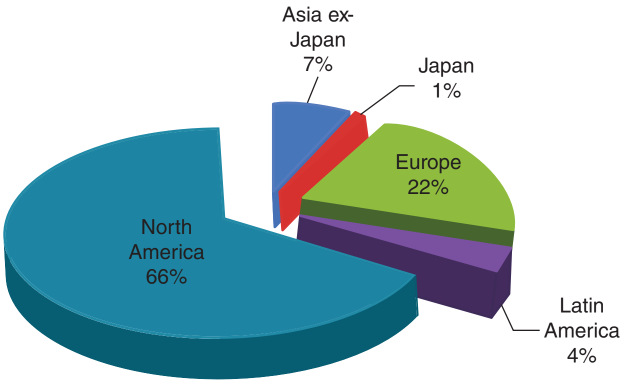


Figure 1.3 Geographical locations of hedge fund
Source: Eurekahedge

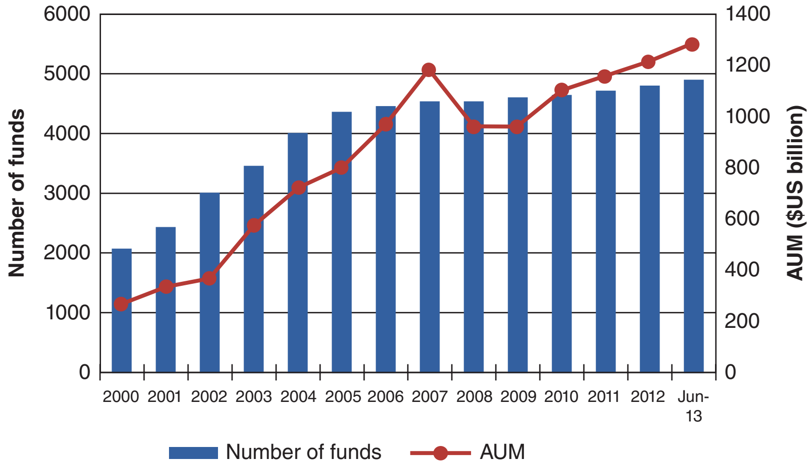


Figure 1.4 Growth of the North American hedge fund industry since 2000
Source: Eurekahedge

1.3.1 North America

Despite periods of high volatility and market swings, North American hedge funds have consistently posted record returns since reaching their lowest point in early 2009. The total size of the industry at the end of 2012 was estimated at $1.15 trillion, managed by over 4800 funds (see Figure 1.4). This is a clear indication of the confidence investors began to show in North American funds after the fallout from the global financial crisis of 2008 when billions of dollars were redeemed and funds suffered massive performance-based losses. Since then, hedge fund managers have provided significant protection against market downturns as well as addressing investors' concerns over counterparty risk by engaging multiple prime brokers instead of the usual singular relationship. Moreover, managers have increased redemption frequencies allowing investors better access to their capital, allowed for more transparency across investment strategies and implemented more stringent risk management controls. Such changes, together with a much improved outlook on the US economy and the introduction of *quantitative easing*,[11] has led to increased investor confidence and substantial asset flows into North American hedge funds which looks set to continue well into 2014.

[11] *Quantitative easing* is a monetary policy that has been employed by the US, the UK and the Eurozone, especially, since the financial crisis of 2008. When a country's interest rate is either at, or close to zero, normal expansionary monetary policy fails so the central bank creates new money which it uses to buy government bonds and increase the money supply and excess reserves of the banking system. A further lowering of interest rates follows and it is anticipated that this will lead to a stimulus in the economy.

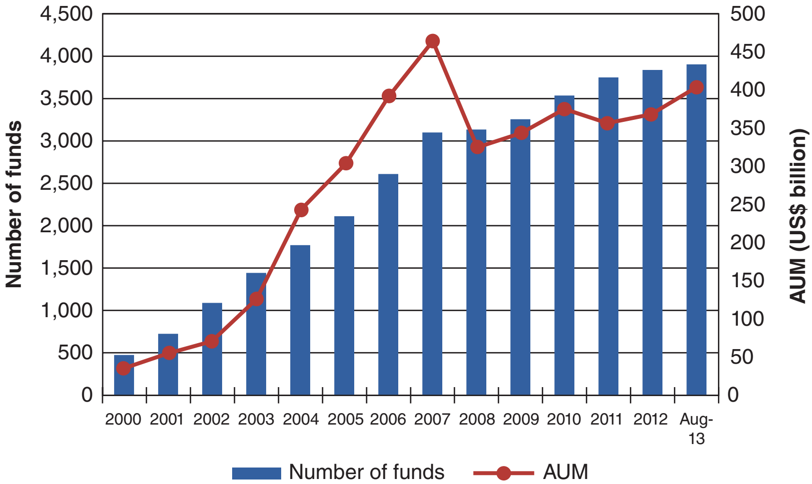


Figure 1.5 Growth of the European hedge fund industry since 2000
Source: Eurekahedge

1.3.2 Europe

The rapid growth of the European hedge fund industry over the first seven years of the last decade was eventually slowed by the onset of the financial downturn in 2008. As with North American hedge funds, the European sector experienced huge losses and increased pressure for redemptions from investors which continued until early 2009 when the global economy began to see a recovery (see Figure 1.5).

The European region has shown some interesting trends with regards to fund launches since the market began to rebound in 2009. However, in 2010, the fortunes of European hedge funds were hit again as a result of the on-going European sovereign debt crisis. Although attrition rates have been relatively high, launches have gained strength on the back of the new UCITS III regulation.[12] The popularity of UCITS III has seen the launch of many new start-ups seeking investment capital in the increasingly competitive hedge fund arena. Many new hedge fund launches have suffered from the investment bias towards allocating to much better-known and larger-based hedge fund names. However, it is anticipated that this trend is likely to change as a result of the increased diversification offered by European hedge funds and a new regulatory environment over the coming years.

[12] UCITS III is the *Undertakings for Collective Investment in Transferable Securities*, an EU investment regulation for the creation and distribution of pooled investment funds, such as mutual funds and hedge funds (see Section 1.5.1).

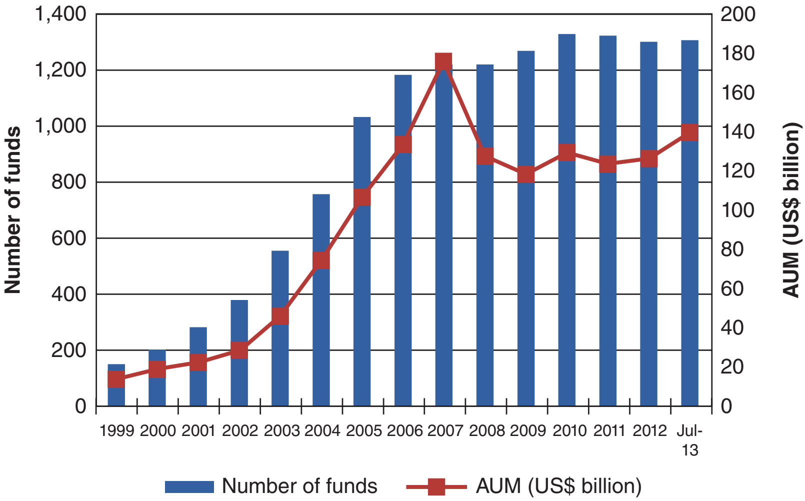


Figure 1.6 Growth of the Asian hedge fund industry since 1999
Source: Eurekahedge

1.3.3 Asia

Like the European hedge fund sector, Asian funds saw tremendous growth over the last decade up until the slowdown during the financial crisis starting in 2008. After the second half of 2009, the Asian sector has shown a steady improvement, with over 1300 active funds, however, the sector has not seen the growth experienced pre-2008. This is mainly due to speculation that the Asia markets may suffer from a possible double-dip recession as a result of the debt contagion passing from Europe (see Figure 1.6).

In the Asian sector, hedge fund managers have struggled to generate asset flows and this, together with the highly volatile markets, has led investors to be extremely cautious about investing in the current climate. However, the desire for Asian governments to attract global hedge fund managers to the region, reductions in hedge fund set-up costs compared to other western locations, the availability of a growing range of financial products and the easing of access and market restrictions in regions such as China and India, should see an increased growth in the sector through 2014 and beyond.

1.4 SPECIALIST INVESTMENT TECHNIQUES

1.4.1 Short Selling

A *short sale* is the sale of a security that a seller does not own or that is completed by delivery of a borrowed security. The short seller borrows the securities from a prime broker in return for paying a daily fee, and promises to replace the borrowed securities at some point in the future. The transaction requires the prime broker to borrow the shares from a securities lender and make delivery on behalf of the short seller. Prime brokers can borrow securities from custodians who hold large institutional investments, e.g. mutual and pension funds or from their own proprietary trading accounts. The cash from the transaction is held in an *escrow account*[13] until the short seller is in a position to replace the borrowed shares (or they are called back by the lender). Since the short seller borrows the securities from the prime broker and has a future commitment to replace them, collateral must be posted in the form of cash, securities or other financial assets. The collateral, in addition to the fee for borrowing the securities, provides the prime broker with additional income in the form of interest until the shares are returned.[14] In addition, the short seller must cover any dividends paid on the shares during the period of the loan and in the case of any stock splits, e.g. two-for-one, the short seller must pay back twice as many shares.

The eventual buyer of the shares from the short seller is usually unaware that it is a short sale so the seller must make arrangements to cover the delivery obligations. The shares are transferred to the buyer with full legal ownership, including voting rights which can pose a severe problem for the short seller if the prime broker requires the securities back (or *called-away*), for example if the original securities lender requires them for a company shareholder meeting. Although this rarely happens in practice short selling does carry a great deal of risk, especially if the shares are held over a long period of time and the stock fails to decline as expected causing them to have to post further margin and eventually forcing the short seller to close out their position at a significant loss. However, when stock prices fall, short sellers make a profit from the short sale, and also between 60–90% of the interest income charged by the prime broker on the cash deposit (i.e. the *short rebate*).

It is often the case that hedge funds do not disclose the names of companies they are selling short to investors for fear of a *short-squeeze*. Unexpected news on short selling activity can cause share prices to suddenly rise due to potential price manipulation through long investors buying additional shares or forcing securities lenders to recall loaned shares. In this case, short sellers' demand for stocks to cover their short positions can cause a mismatch between the availability of shares and thus drive prices up further. To avoid short-squeezes, hedge funds employing short selling only normally invest in large cap companies which have a greater amount of liquidity and volume of shares available from prime brokers. In the US, hedge funds are only allowed to engage

[13] An *escrow account* is an account set up by a broker for the purpose of holding funds on behalf of the client until completion of a transaction.

[14] Borrowing money to purchase securities is generally known as *buying on margin*. It is usually necessary for a hedge fund to open a margin account with a prime broker and maintain the margin with available cash reserves when market prices move adversely in order to meet a *margin call*.

in short sales with those securities whose recent price change was an upward movement.[15] Such restrictions are used to prevent hedge funds investing in stocks that are already declining so as to avoid the possibility of sending the market into free fall. However, short sellers are often thought of as providing efficient price discovery as well as market depth and liquidity. It is important to investors that they are confident that prices represent fair value and that they can get easy access to liquid markets in which they can readily convert shares into cash. Hedge funds through short selling provide this level of confidence by forcing down overvalued stocks and generating liquidity within the markets.

1.4.2 Leverage

Leverage is using borrowed cash, or a margin account, to increase the purchasing power and exposure to a security (or investment) with the aim of generating higher returns. Financial instruments, such as options, swaps and futures (i.e. derivatives) are also used to create leverage. A premium is paid to purchase a derivative in the underlying asset which gives them various rights and obligations in the future. This premium is far less than the outright price of the underlying asset and thus allows investors to buy an economic exposure to considerably more of the asset than they could otherwise.

Although generally misunderstood, leverage is an extremely widespread investment technique, especially in the hedge fund industry. A great deal of confusion often arises from the various definitions and measurements of leverage. In terms of hedge fund leverage, the debt-to-equity ratio or percentage is often the preferred indicator. For example, if a hedge fund has \$50 million equity capital and borrows an additional \$100 million, the fund has a total of \$150 million in assets and a leverage of $2 \times equity$ or 67% (=100m/150m). Leverage ratios are typically higher than traditional investments and generally more difficult to measure due to the sophisticated use of certain financial instruments and strategies. Since adding leverage to an investment inherently increases risk, investors often equate a highly leveraged hedge fund with a high risk investment. However, this is not normally the case since hedge funds often use leverage to offset various positions in order to reduce the risk on their portfolios.[16] For this reason, it is not advisable for investors to solely rely on leverage ratios as proxies for hedge fund risk. It makes more sense to correctly

[15] The *up-tick* rule was introduced in the US by the Securities and Exchange Commission (SEC) in 1938 to restrict the short selling of stocks unless there was an upward movement in the price. The restriction was lifted in 2007, but there has since been a growing debate on reinstating the ruling to prevent the potential for market manipulation through short selling (see Section 1.5.3).

[16] The amount of leverage that hedge funds can take on is usually limited by margin supplied by the prime broker and on certain restrictions set by regulators or other organisations. In circumstances where the amount of leverage rises above a certain limit, the lender can take possession of the hedge fund investments, sell them and use the proceeds to offset any losses on the debt financing.

analyse the nature of the strategy in more detail before making a decision on the riskiness of the hedge fund.

1.4.3 Liquidity

Although hedge funds can generate abnormal returns by exploiting the value from investing in *illiquid*[17] assets, there is always a need to access market liquidity. Liquidity is the degree with which an asset can be bought or sold without adversely affecting the market price or value of the asset.[18] Liquidity plays a critical role in the financial markets providing investors with an efficient mechanism to rapidly convert assets into cash. During the recent financial turmoil, hedge funds experienced an unprecedented number of requests from investors to withdraw their capital creating a serious liquidity problem.

The global credit crisis originated from a growing bubble in the US real estate market which eventually burst in 2008. This led to an overwhelming default of mortgages linked to subprime debt to which financial institutions reacted by tightening credit facilities, selling off bad debts at huge losses and pursuing fast foreclosures on delinquent mortgages. A liquidity crisis followed in the credit markets and banks became increasingly reluctant to lend to one another causing risk premiums on debt to soar and credit to become ever scarcer and more costly. The global financial markets went into meltdown as a continuing spiral of worsening liquidity ensued.

Hedge funds that had assets linked to the subprime debt disaster and other related securities suffered huge losses. Problems were amplified further when investors tried to withdraw capital from their funds and it became apparent that there was a liquidity mismatch between assets and liabilities. When the credit markets froze, hedge fund managers were unable to source sufficient capital to meet investor redemption requirements. This forced managers to restructure their liquidity terms and impose further *gate provisions*,[19] increase the use of *side-pocketing*[20] and enforce *lock-ups*.[21] Not only did this negatively affect investor relations but this was further damaged by the selective and insufficient disclosure of performance being made by hedge fund managers. For example, many managers were seen to be reporting side-pocket performance only to investors while relaying much better liquid performance in publications with

[17] *Illiquid* assets include low volume traded stocks, real estate and other capital holdings.

[18] A highly liquid market can also be considered a *deep* market.

[19] *Gated provisions* are a restriction on the amount of capital that can be withdrawn from a fund during a redemption period. Such provisions are subject to management discretion and normally referred to in the hedge fund prospectus.

[20] A *side-pocket* (or *designated investment*) is an account used by hedge funds to separate illiquid assets from more liquid ones. Holding illiquid assets in a hedge fund can cause a great deal of complexity when investors try to withdraw their capital.

[21] A *lock-up* is a period of time designated by the hedge fund manager in which an investor may not withdraw any investment in a fund.

only a passing note of disclosure about the exclusion of side-pockets. In the case of lock-ups, there exists a clear conflict of interest between locking up investors' capital and continuing to charge management fees. Investors have since argued that it would be more acceptable for gate provisions and the issue of involuntary side-pockets to be tied to deferrals or a reduction in management fees until the fund returns to an appropriate liquid position. The aftermath of the financial crisis has clearly highlighted many of the shortcomings of the hedge fund industry and heightened the debate over the need for increased regulation and monitoring. Nevertheless, it has since been widely accepted that hedge funds played only a small part in the global financial collapse and suffered at the hands of a highly regulated banking system. Indeed, many prominent institutional and economic bodies argue that their very presence provides greater market liquidity and improved price efficiency whilst aiding in the global integration of the financial markets.

1.5 NEW DEVELOPMENTS FOR HEDGE FUNDS

1.5.1 UCITS III Hedge Funds

One of the major developments in the hedge fund industry over the past several years has been the exceptional growth in UCITS III hedge funds in comparison to the global industry. As of the end of 2012, the UCITS hedge fund industry stood at an estimated $215 billion managed by over 900 individual funds (see Figure 1.7). UCITS is a set of directives developed by the EU member

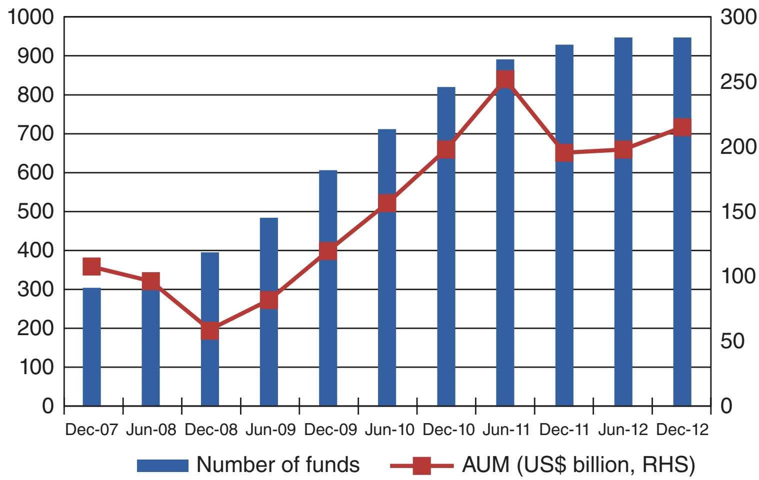


Figure 1.7 Growth in number of UCITS III hedge funds since December 2007
Source: Eurekahedge

states to allow cross-border investments. The aim of the directive is to improve the financial opportunities offered to UCITS-compliant hedge fund managers whilst addressing the needs of investors in terms of effective risk management procedures, increased transparency and liquidity, especially in light of the recent financial crisis.

The original version of the directive was introduced in 1985 with the goal of establishing a common legal framework for open-ended funds investing in transferable securities set up in any EU member state. That is, to develop a pan-European market in collective investment schemes. Unfortunately, the framework suffered from many obstacles, such as the extent of different marketing rules and taxation allowed across member states. Not until December 2001 was a directive formally adopted under the UCITS III banner which has since undergone several further amendments with a view to including the use of additional asset classes (e.g. hedge fund indices) and a more diverse range of derivative products (see Figure 1.8). Such inclusions have allowed UCITS III funds to pursue a number of different investment possibilities, such as absolute return strategies, in ways that were simply not possible under previous UCITS frameworks.

The increased number of eligible asset classes and available use of derivatives has led to a greater number of multi-strategy funds being launched in the UCITS III sector in comparison to that of the European multi-strategy industry. Despite

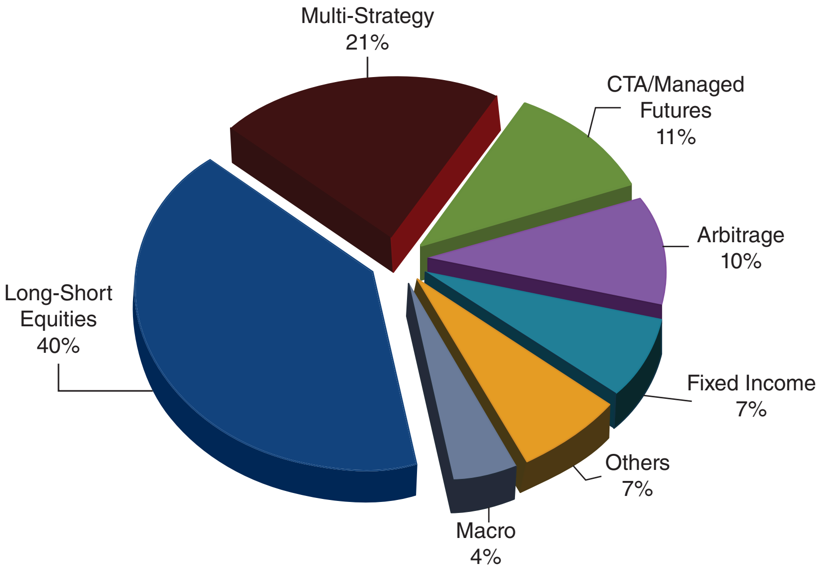


Figure 1.8 UCITS III hedge funds by investment strategy
Source: Eurekahedge

this, however, almost half of UCITS III funds over the last two years have adopted the Long-Short Equity strategy for several reasons:

1. Long-Short Equity is by far the largest global hedge fund strategy and therefore those existing managers launching new UCITS III vehicles would naturally prefer this strategy.
2. Operating the Long-Short Equity strategy under the UCITS III framework is relatively straightforward.
3. The simplicity of the Long-Short Equity strategy lends itself well to marketing and liaising with retail investors who may not have the knowledge and understanding of the markets like a typical hedge fund client.

Despite the strong link in the use of the Long-Short Equity strategy, the regulatory constraints within the UCITS III framework mean there are very few similarities elsewhere across industry sectors. For example, there are very few event-driven UCITS III hedge funds and practically no distressed debt-based funds primarily due to the liquidity restrictions placed on UCITS III-compliancy. Nevertheless, a major advantage of UCITS III funds is the ability of managers to utilise their experience in a proven investment strategy whilst offering potential investors the added incentive of investing in a regulated market. Some of the key features of UCITS III hedge fund regulation and fund structure include the following:

1. Only investment in liquid securities is permitted, i.e. those that can be sold within 14 days without substantial loss of value.
2. Funds cannot have exposure to more than 10% in one stock.
3. Managers can utilise leverage up to 100% of the *Net Asset Value*[22] (NAV) of the fund.
4. Managers can employ shorting techniques through the use of derivatives.
5. The ease of marketing such funds across the EU and registering them in member states.
6. The implementation of effective risk management procedures and processes.
7. Funds must be domiciled in an EU member state as opposed to offshore locations.

The development of UCITS III funds has also opened up the sector to new sources of capital, for example, from retail investors wishing to make use of the alternative investment market but with the assurance of stricter regulatory controls. In addition, improved redemption rules and transparency have helped in building investor confidence, especially after the much debated issue surrounding the use of gated provisions that stopped investors withdrawing large amounts of capital from their funds during the period of huge losses following

[22] Net Asset Value (NAV) is used to put a value on a hedge fund and is the total of all the hedge fund assets minus all the hedge fund liabilities.

the financial crisis in 2008. The release of UCITS IV in mid-2011 and the development of other European directives, such as the new EU Passport, which will give hedge funds marketing rights throughout the EU, will broaden the investment appeal of UCITS-compliant funds even further. UCITS IV provides investors with more transparency, facilitates cross-border hedge fund distributions, reduces costs and achieves regulatory alignment.

1.5.2 The European Passport

In November 2010 the European Union passed a new set of laws governing the use and regulation of the alternative investment industry, named the Alternative Investment Directive (AID). The AID aims to provide hedge funds (and private equity funds) with a so-called *passport* to allow funds that meet EU standards access to all EU markets. The passport gives hedge funds the opportunity to market to investors throughout the EU with only a single authorisation. In addition, the directive will subject hedge funds to increased supervision, regulation and transparency providing pan-European investors with the confidence to invest and operate in a stable European financial market.

The newly formed Paris-based European Securities and Markets Authority (ESMA) will act has the EU financial supervisory authority and issue passports, especially to non-EU funds that wish to operate in the EU markets under a single authorisation. The ESMA will also demand non-EU funds to grant the same rights that their funds will enjoy in the European markets. However, the controversial passport scheme will not come into effect for EU funds until late 2013, and even later for non-EU funds, so the established framework that allows each EU country to decide which funds they will allow access to their market will remain in place until then.

1.5.3 Restrictions on Short Selling

Short selling can result in unlimited losses if the hedge fund incorrectly anticipates the direction of movement of share prices. Moreover, short selling can also be used to manipulate market prices. It has been argued for some time that hedge funds can engage in collective short selling to create an imbalance in supply and force down the price of a security. During the recent financial turmoil and substantial falls in stock prices, hedge funds were often accused of short selling to exacerbate and profit from the declining markets. During 2008 and 2009, regulators announced several actions to protect against abusive short selling and to make short sale information more readily available to the public.

One of the main methods of market abuse was the use of *naked* short sales, i.e. the activity of selling short without having borrowed or arranged to borrow the securities to make delivery to the buyer. Such a *failure-to-deliver* is a gross violation of ethical market practice and something the regulators were determined to address. New temporary rulings forced prime brokers to first ascertain, before undertaking a short sale transaction, if the securities were available for short selling or could be borrowed against delivery. Market participants were also required to provide detailed information on short sale activity and their overall short positions. Although the rulings curbed short sale abuses during the financial crisis of 2008 many hedge fund managers have since argued that such regulation hinders the efficient workings of the financial markets and causes a negative effect on liquidity. Restrictions and other regulation on short sales is a contentious area of debate amongst market professionals and regulators and under continual review. Nevertheless, regulators are keen to address issues associated with short selling and the provision of detailed information on short sale activity for public disclosure and may certainly impose further restrictions on the practice in the near future.

In this chapter we have provided an introduction to the concept of a hedge fund, how they are structured and the key players within such an investment vehicle. Each of the major markets within the global hedge fund industry has been reviewed with focus on the current financial crisis and how hedge funds have performed over this period. In fact, it has been publicly stated that hedge funds have played only a minimal part in the global financial collapse and have instead suffered at the hands of a highly-regulated financial system. Nevertheless, some of the specialist investment techniques employed by hedge fund managers have since come under increased scrutiny and regulatory pressures.

2

Hedge Fund Data Sources

Obtaining accurate, reliable and timely data on hedge funds is of extreme importance to a manager or analyst wishing to measure monitor and assess the returns and performance of such investment opportunities. The need for consistent and robust hedge fund indices, as well as trustworthy benchmarks for the industry, is also of extreme value and a necessity when trying to ascertain a clear representation of a hedge funds track record.

Chapter 2 reviews a variety of prominent commercial hedge fund databases and the variety of indices and benchmarks they produce. In addition, the many pitfalls and problems that need to be fully understood when interpreting and using such summary statistics are discussed. In particular, the inherent heterogeneity and lack of representativeness within the hedge fund universe is highlighted as a major concern in the industry. Moreover, some of the most innovative products developed in the market in order to overcome some of these constraints are reviewed.

2.1 HEDGE FUND DATABASES

Many hedge fund managers provide, on a voluntary basis, monthly hedge fund performance data to a variety of commercial databases. These hedge fund databases collect, assimilate and produce informative reports, indices and benchmarks based on this data for potential investors, consultants, analysts and academics involved in investment and research on hedge funds. Such data allows the construction and publication of numerous non-investable and investable indices that purport to give an indication of the state of the hedge fund industry and act as important benchmarking indicators. However, due to the voluntary nature of a fund manager's requirement to supply informative monthly data, the indices produced by these vendors can be misleading and contain several biases and anomalies. Moreover, a complete record of every single hedge fund in the industry is simply not available; relevant information comes only from samples of the hedge fund universe. The quality and quantity of such data vary between vendors and an investor is left to their own judgement in accepting one set of hedge fund performance statistics over another. Over the past several years hedge fund managers have chosen to be much more selective in the databases to which they report. For this reason, it is important to note that any given database provides only a sample of the entire hedge fund universe, requiring investors to access multiple data sources to get a clearer understanding of the investment opportunities and the state of the hedge fund industry.

2.2 MAJOR HEDGE FUND INDICES

In traditional finance, the use of indices is a useful investment tool for managing the exposure of a portfolio of investments to *market risk*, i.e. a so-called *passively* managed investment style. In the hedge fund world, where performance is generally driven by a manager's skill and expertise, reflecting an *active* management style, the use of hedge fund indices is more surprising, since the concept of an index is commonly associated with the notion of passive management. However, many of the commercial databases, and a selection of traditional index providers, e.g. FTSE and Dow Jones, have developed and published a range of hedge fund indices and relevant benchmarks.[1] The way such indices are designed and constructed varies considerably amongst providers and should be fully understood before accepting the index has a valued industry benchmark or performance measurement. Many of the index providers purport to have the best index methodology and construction process with strict inclusion criteria and thorough due diligence procedures. Nevertheless, many investors, as well as the industry itself, are wary of considering one index better than another and approach the hedge fund index and benchmark arena with caution.

2.2.1 Non-Investable and Investable Indices

The hedge fund industry is highly heterogeneous making the construction of a satisfactory performance index that comprises the available hedge fund universe extremely difficult. Non-investable hedge fund indices try, at best, to represent the performance of a sample of the hedge fund universe taken from a relevant database. However, such databases have diverse selection criteria and methods of index construction leading to many differing published indices. Although aiming to be representative of the hedge fund universe, non-investable indices suffer from many unavoidable *biases* (see Section 2.3). A further difficulty associated with the heterogeneity of hedge funds is the classification of investment styles. With around 10,000 active hedge funds in the industry, determining each manager's investment style is virtually impossible. Some index providers have developed their own classification system that attempts to capture a high level of homogeneity within each investment group and a subsequent level of heterogeneity between individual groups. Unfortunately, many hedge fund managers do not report their investment style correctly or often change them without giving prior notice to the database vendor. Clearly

[1] Over the past decade, many hedge fund databases and index providers have merged, been acquired by larger firms, or developed into *boutiques* offering specialist hedge fund services and consultancy.

then, hedge fund indices are subject to a greater lack of representativeness than traditional indices. Such a problem goes beyond an insufficient classification of investment styles. Instead, it concerns the actual managers themselves who have a great deal of freedom and choice at their disposal.

By early 2000, many index providers had launched a series of investable indices offering a low-cost investment opportunity to gain exposure to the hedge fund industry. An index is investable when the investors are able to replicate the index by obtaining and maintaining a certain level of *tracking error* (see Section 2.4.1). Generally based on platforms of separate *managed accounts*,[2] this new generation of indices has been able to provide investors with improved liquidity and a low-cost method of gaining access to the hedge fund world. In addition, the composition, construction methodology and management principles are overseen by an independent committee and disclosed to the public giving increased transparency. To create an investable index, the index provider selects a range of hedge funds and develops structured products or derivative instruments that deliver the performance of the investable index. To make them investable, each hedge fund agrees to accept investments on the terms given by the index provider. When investors buy an investable index the provider makes the investments in the underlying hedge funds, making such an index similar in many ways to a fund of hedge fund (FoHF)[3] investment. In fact, some refer to investable indices as merely FoHFs in disguise or with additional constraints. However, by their very construction, investable hedge fund indices are unable to represent the hedge funds universe. Such indices cannot represent the *open* funds universe, since they are not composed of funds belonging to this subset of the complete universe. They contain many *partially* closed hedge funds, i.e. funds that accept new investments only from investors that have reserved capacity. Hedge funds generally do have capacity issues as certain strategies only work well within certain limits of investment capital. Indeed, many hedge fund managers refuse further capital into the fund after they have reached a maximum level of AuM and assign them as *closed* funds. As a result, it is very difficult for hedge fund indices to remain investable when the composite hedge funds have closed their doors to new investors. Most index providers argue that to be a truly representative index that acts as a gauge for hedge fund performance, both open and closed funds should be included in the hedge fund index. The trade-off, therefore, is between having as broad a representation as possible of hedge fund performance against having a smaller sample of hedge fund managers representing the performance accessible through investment.

---

[2] *Managed accounts* are a rapidly-growing, fee-based investment management product for High Net Worth (HNW) individuals. Such accounts allow access to professional money managers, high degrees of customisation and greater tax efficiencies. They are also said to provide the added benefits of greater TLC, i.e. Transparency, Liquidity and Control.

[3] A *fund of hedge fund* (FoHF) is a common investment vehicle for inexperienced investors or those that have limited exposure to the alternative investment market.

By the end of 2006, hedge fund *replication* aimed to eliminate many of the problems associated with hedge fund indices. Instead of accessing the performance of hedge funds they take a statistical approach to the analysis of historic hedge fund returns, and use this to construct a model of how hedge fund returns react to movements in a range of investable financial instruments. This model is then used to construct an investable portfolio of those assets. This makes the index investable, and in principle they can be as representative as the hedge fund database from which they were constructed. As the hedge fund industry becomes increasingly diversified offering greater opportunity for investors to gain exposure to hedge fund returns without direct involvement, the growth in such indices will surely increase.

2.2.2 Dow Jones Credit Suisse Hedge Fund Indices (www.hedgeindex.com)

The Credit Suisse/Tremont Hedge Fund Indices were rebranded in 2010 after Credit Suisse joined forces with Dow Jones Indices to provide the flagship Dow Jones Credit Suisse (DJCS) Hedge Fund Index. Dow Jones is responsible for the calculation, distribution and marketing of the indices whilst Credit Suisse affiliates continue to manage the financial instruments associated with them.

The DJCS Hedge Fund (or *Broad*) Index is one of the industry's most respected indices which tracks approximately 9000 funds from the proprietary Credit Suisse database and includes 400 hedge funds as of July 2010. The index comprises only hedge funds with a minimum of \$50m AuM[4] and audited accounts,[5] although hedge funds with an AuM of more than \$500m and a track record of less than a year may be considered under special circumstances. In terms of index participation, AuM does not include managed accounts and reflects only the assets of the fund, not the AuM of the investment company. For index inclusion, fund managers must report performance returns (or NAVs) and AuM on a monthly basis. The index is asset-weighted (see Box 2.1) and broadly diversified across 10 style-based investment strategies (see Figure 2.1) which seeks to be representative of the entire hedge fund universe. The DJCS Hedge Fund Index construction is based on a transparent and rule-based selection process. The methodology analyses the percentage of assets invested in each subcategory and selects funds for the index based on those percentages, matching the shape of the index to the shape of the universe. Fund *weight caps*

[4] The amount of AuM can sometimes be difficult to determine, since hedge fund managers combine managed accounts and onshore/offshore vehicles, and also use different amounts of leverage, either through margining or by short selling.

[5] A hedge fund must have a minimum one-year track record to have audited financial accounts.

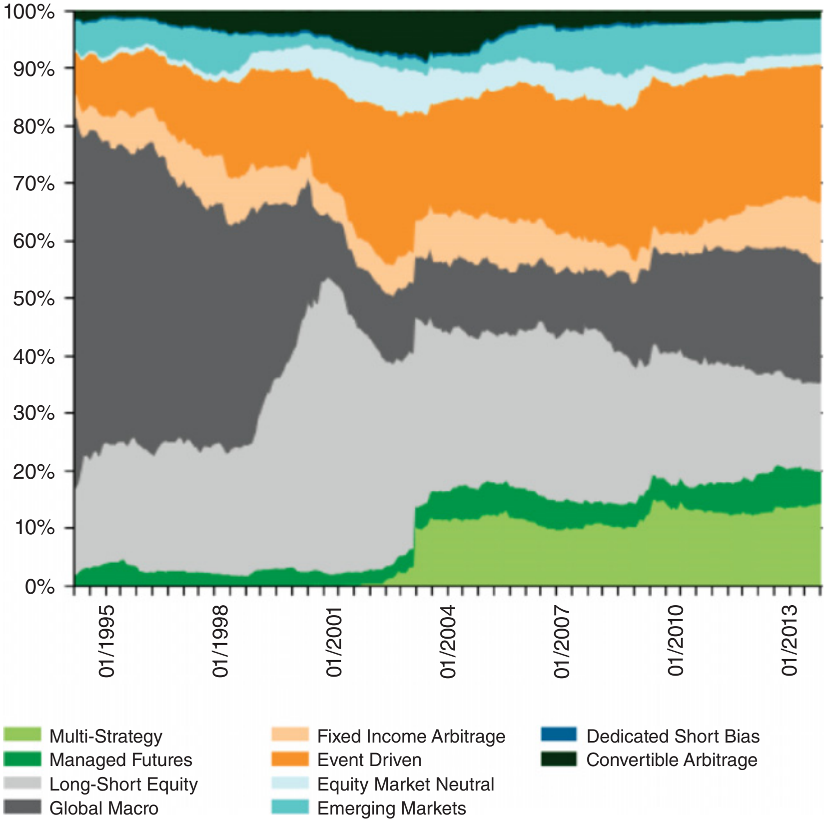


Figure 2.1 Historical asset weights by investment strategy for the DJCS Hedge Fund Index
Source: Hedge Index

can be applied to enhance diversification and limit concentration risk. The index is calculated and rebalanced monthly and funds are reselected on a quarterly basis as required.

Box 2.1 Asset Weighting Schemes

When measuring the performance of a portfolio of hedge funds or an average of a group of funds, it is necessary to assign a particular weight to each of the individual funds. There are three major weighting schemes used in the hedge fund industry, namely *equal*, *asset* and *arbitrary* weightings.

Equal

If there are N hedge funds in the group of funds, then, each fund return has an equal weight, w_i in the average that is given by:

$$w_i = \frac{1}{N}$$

The average is a measure of the average behaviour of the fund returns irrespective of the amount of AuM (or *market capitalisation*) of each hedge fund.

Asset

Each hedge fund return is weighted with respect to the amount of AuM in proportion to the total assets managed within the group of funds, i.e. dollar-weighted averages. If a particular fund i has AuM denoted by A_i, then the weight of fund i in the average is given by:

$$w_i = \frac{A_i}{\sum_{i=1}^{N} A_i}$$

Arbitrary

Each hedge fund return is given an arbitrary weight w_i within the average which can be changed over time. However, the total of all the arbitrary weights must always sum to 100%.

Suppose we have a number of hedge funds, N in a group where the return on fund i is denoted by r_i and respective weight w_i, then a performance measurement (or index) for the group can be determined as the weighted average of the individual hedge fund returns, such that:

$$r_{index} = \sum_{i=1}^{N} w_i r_i$$

The index family currently consists of 17 indices, including a range of geographical and strategy-specific hedge fund indices. The current family includes:

- The *DJCS AllHedge Index* was formerly the CSFB/Tremont Sector Invest Indices, is a diversified investable index comprised of an aggregate of all 10 DJCS AllHedge Strategy Indices and asset-weighted based on the sector weights of the Broad Index. The AllHedge Index was launched in October 2007, and any performance of the index predating October 2007 is simulated

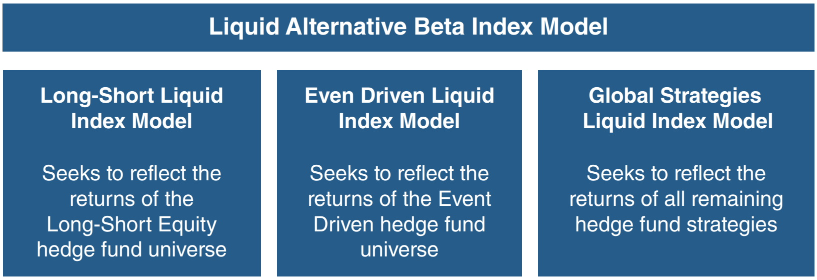


Figure 2.2 The Liquid Alternative Beta Index Model
Source: Hedge Index

from returns on the underlying AllHedge Strategy Indices as of October 2004. AllHedge seeks to represent the investable hedge fund universe and encompasses around 80 funds as of November 2013,

- The *DJCS Blue Chip Hedge Fund Index*, formerly the CSFB/Tremont Investable Hedge Fund Index, is an investable index made up of the 60 largest hedge funds from the 10 style-based sectors compromising the Broad Index, and
- The *DJCS LEA Hedge Fund Index* is an emerging market asset-weighted composite[6] index covering the regions of Latin America, Europe, Middle East, Africa and Asia emerging economies. The index was launched in April 2008.

2.2.2.1 Liquid Alternative Betas

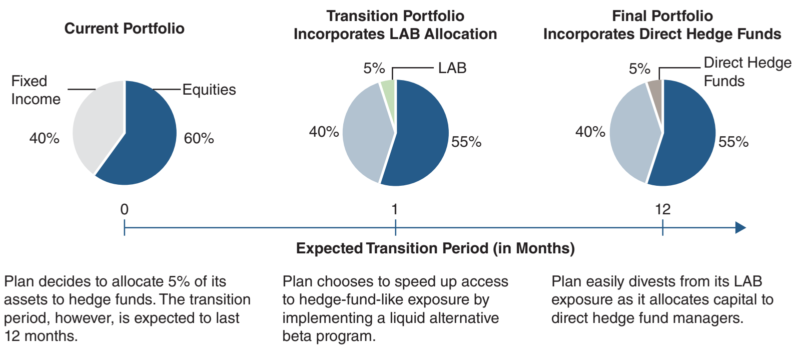
In addition to the above index family, Credit Suisse publishes a series of Liquid Alternative Beta (LAB) indices.[7] LAB indices aim to replicate (or *clone*) the aggregate return characteristics of alternative investments strategies using commonly-traded instruments with high liquidity. LABs reflect the returns of a dynamic basket of investable market factors selected and weighted so as to approximate the aggregate returns of the universe of hedge funds represented by the family of DJCS hedge fund indices. Such liquid replication strategies seek to provide hedge fund returns without direct hedge fund investment and thus enhance liquidity and eliminate hedge fund *headline*[8] risk (see Box 2.2). The range of LAB indices currently includes the Event Driven Liquid Index, Global Liquid Index, Long-Short Equity Liquid Index, Merger Liquid Index and Managed Futures Liquid Index as shown in the schematic in Figure 2.2.

[6] A composite index consists of individual hedge funds that cover a range of different investment strategies.
[7] Alternative beta refers to alternative systematic risks, i.e. those risks that cannot be diversified away and are compensated through risk premia or the expected rate of return above the risk-free interest rate.
[8] *Headline* risk is the impact that a negative news story could have on the value of an investment.

Box 2.2 A LAB Example

Problem: A pension plan with a traditional 60/40 (equity/fixed income) portfolio is planning a 5% hedge fund allocation. To implement the program, the plan starts due diligence on several hedge fund managers, a lengthy process that may delay capital deployment. If the funds remain in cash or short maturity fixed income, expected returns could be negatively impacted.

Solution: To better manage the transition and speed up the exposure of the plan to potential hedge-fund like returns, the plan makes an allocation to a liquid alternative beta strategy with daily liquidity.* The risk/return profile of this interim allocation is expected to provide a reasonably close match to that of the direct hedge funds in which it plans to invest. The plan can then draw down its replication exposure gradually to reallocate to the selected hedge funds as it completes the due diligence process (see diagram below).



For illustrative purposes only. Does not represent an actual investment or the actual performance of any product or portfolio.

*The liquid alternative beta Long-Short Equity data in this example is represented by the CS Long-Short Equity Liquid Index.

Source: Credit Suisse Asset Management

The three main components that drive hedge fund performance can be identified as:

1. Traditional beta,
2. Alternative beta, and
3. Alpha.

Table 2.1 Traditional and alternative beta factors driving aggregate hedge fund returns and their respective proxies

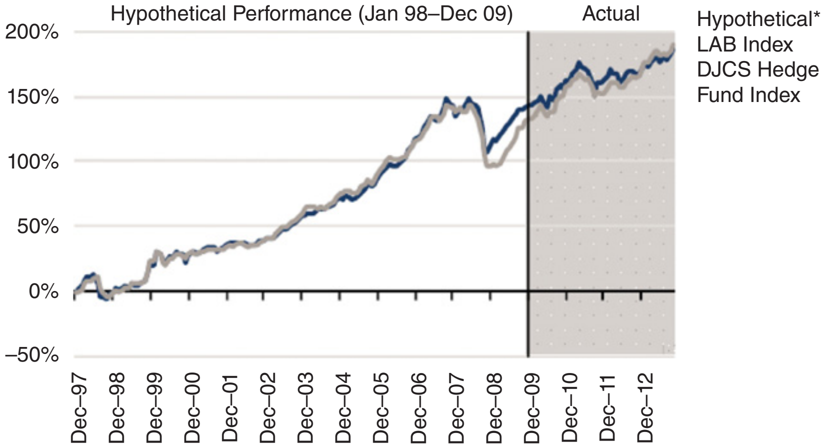
	Strategy	Factors	Proxies
Traditional Beta Examples	Long-Short equity	US large cap	S&P 500 index
	Long-Short equity	US small cap	Russell 2000 index
	Event driven	High yield fixed income	IBOX high yield liquid index
Alternative Beta Examples	Global macro	Currency carry	Custom currency carry proxy: long high-yielding currencies; short low-yielding currencies
	Long-Short equity	Equity momentum	Custom trend-following proxy: long well-performing companies; short poorly performing companies
	Event driven	Merger arbitrage	Custom M&A proxy: long companies being acquired; short companies that are acquirers

Source: Credit Suisse

Traditional beta returns are based on long-only investment strategies and have exposure to traditional market factors, such as equity and credit risk that act as a *proxy*[9] for well-known benchmark indices, e.g. the S&P 500 and Russell 2000 indices. Alternative beta involve other factors, such as currency carry and equity momentum, that can be captured using various investment strategies and proxied through *systematic* trading.[10] Alpha is attributable to the skill and expertise of the hedge fund manager and is often difficult to capture. Fung and Hsieh (2004), pioneers in the field of hedge fund replication, have shown traditional and alternative beta to be the largest contributors to aggregate hedge fund returns. For this reason, one of the first steps to hedge fund replication is to identify traditional and alternative beta factors that represent the exposure of individual hedge fund strategies. These factors then have to be represented by proxies. These turn out to be fairly straightforward in terms of traditional betas (e.g. S&P 500); however, custom proxies have to be developed for many of the alternative beta factors (see Table 2.1).

[9] A *proxy* is an efficient approximation for another investment.

[10] *Systematic* (or *rule-based*) trading involves using an automated system to trade on behalf of the trader. The system makes all trading decisions with respect to the rules set by the trader and the information available at the time. The other type of trading is discretionary; where the trader uses his intelligence and knowledge to make trading decisions with respect to the information available at the time.



*Simulations for the CS LAB Index were used to measure how a portfolio of securities and market indices designed to track hedge fund indices would have performed in the period beginning December 1997. The LAB Index was launched in January 2010 and shown by the vertical black line.

Figure 2.3 Hypothetical performance of the CS LAB Index from Jan 1998 to Dec 2009 and actual historical performance from Jan 2010 through July 2010
Source: Credit Suisse

Once the factors and proxies have been determined it is necessary to ascertain the combination of exposure and weights that best replicate the hedge fund strategies. Investors can gain market exposure to the factors identified by the analysis using a variety of liquid commonly-traded financial instruments, such as index funds, Exchange Traded Funds[11] (ETFs), swaps, listed futures and option contracts. Any factor exposures and their respective weights can be updated periodically based on ongoing quantitative analysis, e.g. factor modelling. Figures 2.3 and 2.4 show the hypothetical and actual performance of the CS LAB index and a schematic of the LAB index construction process, respectively.

2.2.3 Hedge Fund Research (www.hedgefundresearch.com)

Hedge Fund Research (HFR) produces numerous indices of hedge fund performance ranging from industry-aggregate levels down to specific, focused areas of sub-strategy and regional investment. The HFRI Fund Weighted Composite Index, created in 1994, is one of the most widely used standards of global hedge fund performance. The HFRI index is constructed using equally

[11] ETFs are a security that tracks an index, a commodity or a basket of assets like an index fund, but can be traded like a stock on an exchange.

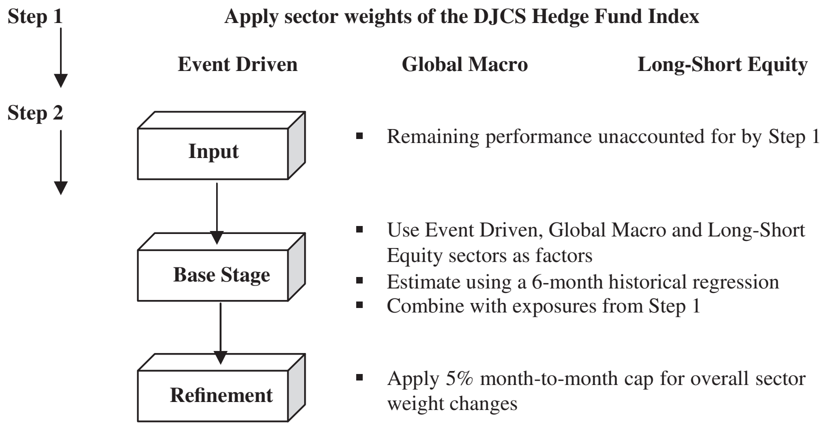


Figure 2.4 A schematic of the CS LAB Index construction process
Source: Credit Suisse

weighted composites of over 2200 single-manager hedge funds reporting to the HFR database. The HFRI FoHF Index is another equally weighted index composite from the HFR database of over 800 FoHFs. Figure 2.5 shows the growth of $1000 since the inception of the HFRI Fund Weighted Composite Index and the HFRI Fund of Funds Composite Index against the S&P 500.

Since 2003, HFR has also produced a range of investable HRFX indices constructed from a variety of quantitative methods, multi-level screening, cluster analysis, Monte-Carlo simulation and optimisation techniques (see Figure 2.6) to ensure that each index is a unique representation of its investment criteria. HFRX Indices are designed to offer full transparency, daily pricing and consistent fund selection, as well as stringent risk management and strict reporting standards.

HFRX Indices use five constituent weighting methodologies and each strategy, sub-strategy and regional investment focus has a corresponding index. The four constituent weighting methodologies are defined as:

1. *HFRX Global Hedge Fund Index* – represents the overall composition of the hedge fund universe comprised of a range of hedge fund strategies; including but not limited to convertible arbitrage, distressed securities, equity hedge, equity market neutral, event driven, macro, merger and relative value arbitrage. The underlying constituents and indices are asset weighted based on the distribution of assets in the hedge fund universe.
2. *HFRX Equally Weighted Strategies Index* – applies an equal weight to all selected constituents.
3. *HFRX Equal Weighted Sub-Strategies Index* – is constructed by equally weighting the sub-strategies included in the HFRX Global Hedge Fund Index.

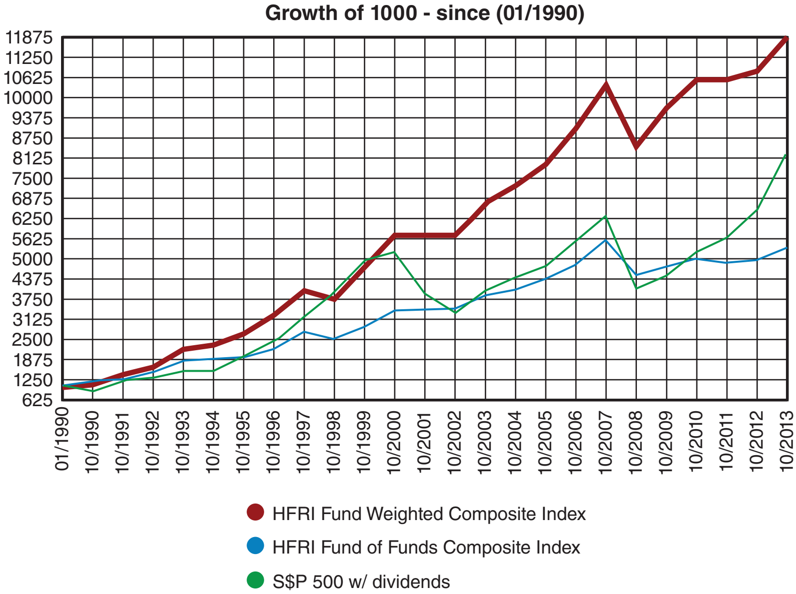


Figure 2.5 The growth of $1000 since inception of the HFRI Fund Weighted Composite Index and Fund of Funds Composite Index against the S&P 500
Source: HFR

Cluster Analysis	HFR screens over 7000 hedge funds to identify those firms with a minimum 24-month track record, at least $50 million AuM, willing to trade on a transparent basis and are open to new investments
↓	
Correlation Analysis	Cluster and correlation analyses are performed to group managers by true strategy categories and to eliminate outliers
↓	
Monte-Carlo Simulation	Monte Carlo Simulation helps determine the adequate number of managers to replicate each strategy
↓	
Due Diligence	Selected managers must provide daily transparency and pass extensive qualitative screening
↓	
Strategy Weighting	Manager investments are then weighted to maximise representation with their group

Figure 2.6 A dynamic, bottom-up approach to HFRX index construction
Source: HFR

Table 2.2 A summary comparison of the HFRI and HFRX indices

Category	HFRI Monthly Indices	HFRX Indices
Created	1994	2003
Weighting	Equally	Index specific
Reporting style	Net of all fees	Net of all fees
Index calculated	Three times a month: Flash Update (5th business day of the month), a Mid Update (15th of the month) and an End Update (1st business day of following month)	Monthly or daily
Index rebalanced	Monthly	Quarterly
Criteria for inclusion	Listing in HFR Database; Reports monthly net of all fees monthly performance and assets in USD	In addition to meeting HFRI criteria, funds must be open to new investment
Minimum asset size and/or track record for fund inclusion	\$50 million or greater than a 12-month track record	\$50 million and 24-month track record
Investable index	No	HFR Asset Management, LLC constructs investable products that track HFRX
Number of funds	Over 2200 in HFRI Fund Weighted Composite Over 500 in HFRI FoHFs Composite	Over 250 in total constituent universe, with over 60 of these in the HFRX Global Hedge Fund Index

Source: HFR

4. *HFRX Absolute Return Index* – incorporates hedge funds that exhibit low volatilities and correlations to standard directional benchmarks of equity market and hedge fund industry performance.
5. *HFRX Market Directional Index* – incorporates hedge funds that exhibit high volatilities and correlations to standard directional benchmarks of equity market and hedge fund industry performance.

Table 2.2 shows a summary comparison of the HFRI and HFRX indices.

Since 2008, HFR has produced several HFRU Indices based on performance benchmarks for hedge funds compliant with established UCITS guidelines. HFRU Indices are representative of the complete universe of UCITS hedge funds, including 4 strategy indices (Equity Hedge, Event Driven, Macro and Relative Value Arbitrage) and an aggregate HFRU Hedge Fund Composite Index. HFRU Indices are published on a daily basis and comprise the most comprehensive benchmarks of UCITS hedge fund performance.

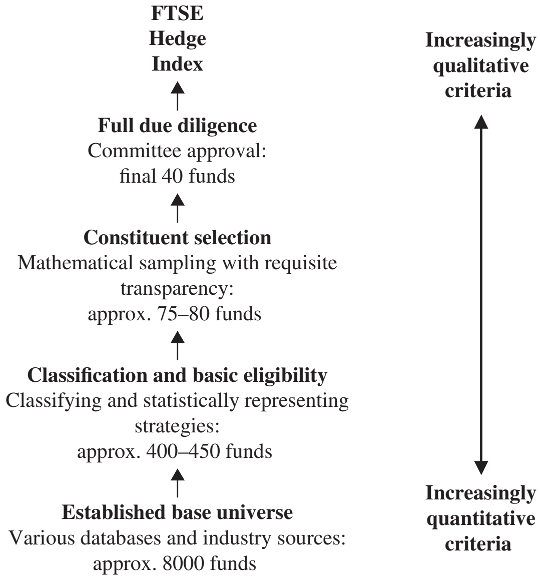


Figure 2.7 FTSE Hedge Index quantitative and qualitative index construction
Source: FTSE Hedge

2.2.4 FTSE Hedge (www.ftse.com)

Since December 1997, FTSE Hedge has provided a series of global indices based on a detailed quantitative and qualitative screening process of around 8000 hedge funds from the FTSE database. The indices give a daily measure of the aggregate risk and return characteristics of a broadly-based global universe of investable hedge funds that allow a sound basis for the creation of liquid and structured investment products within a high degree of transparency. For index inclusion, hedge funds must have a minimum of \$50m AuM and have audited accounts for the past 24 months. FTSE Hedge currently includes the FTSE Hedge Index, eight trading strategy indices (equity hedge, CTA, global macro, merger arbitrage, distressed securities, convertible arbitrage, equity arbitrage and fixed income relative value) and three management style indices (directional, event driven and non-directional).

The construction methodology follows four main stages, namely a classification process, quantitative screening, statistical sampling and a due diligence stage as shown in Figure 2.7. In essence, the funds are classified into strategic categories within the FTSE Hedge Index, which result in around 400–450 funds remaining. These funds are then further screened, by mathematical sampling to ensure transparency, down to 80 funds. Finally, 40 hedge funds are selected, by an independent committee of leading market professionals, for the actual index.

2.2.4.1 FTSE Hedge Momentum Index

Since January 2000, FTSE Hedge also produces the FTSE Hedge Momentum Index, an investment strategy index designed to outperform the underlying FTSE Hedge Index. All constituents of the FTSE Hedge Momentum Index are selected from the FTSE Hedge Index based on strictly defined quality, liquidity and capacity criteria. By under- or over-weighting the constituent funds in terms of whether they show persistent positive return (i.e. momentum), the index can be shown to outperform the FTSE Hedge Index. Indeed, over an eight-year period the FTSE Hedge Momentum Index has returned a 10.1% annualised performance, representing a 3.9% outperformance over and above the FTSE Hedge Index.

2.2.5 Greenwich Alternative Investments (www.greenwichai.com)

First published in 1995, the Greenwich Alternative Investments (GAI) Global Hedge Fund Indices provide more than 20 years of risk and return history that represents both the overall hedge fund universe as well as various constituent groups of hedge funds as defined by their investment strategies. The flagship index, the GAI Global Hedge Fund Index, is designed to reflect the dynamic nature of the hedge fund universe and does not have a fixed set of constituent funds. Instead, GAI attempts to include as many funds as possible, excluding all FoHFs, based on monthly return data. Funds are not excluded on the basis of size, location, or other factors but must have a minimum of three months' track record. Currently around 2000 hedge funds are used to calculate the GAI Global Hedge Fund Index at month-end. Each hedge fund is categorised by investment strategy based on the information supplied by the fund manager and according to the GAI hedge fund strategy definitions. In 2004, four broad Strategy Groups were introduced in the GAI Global Hedge Fund Index, namely Market Neutral Group, Long-Short Equity Group, Directional Trading Group and Specialty Strategies Group. In January 2010, 10 regional indices were introduced as a result of an increasing number of funds focused exclusively on specific geographic regions. The indices were created in two sets, i.e. Developed Markets and Emerging Markets, with each set containing five indices, namely Composite, Global, Asia, Europe and Americas.

2.2.5.1 GAI Investable Indices

The GAI Investable Hedge Fund Indices are an additional series of hedge fund benchmarks designed to represent expected performance of investable hedge funds that are open and considered suitable for institutional investment. The GAI Investable Indices are designed to replicate the performance of the unique strategies of the hedge fund universe (see Figure 2.8).

The GAI Global Hedge Fund Indices form the basis for the construction of the GAI Investable Indices. Asset weights of the Greenwich Investable

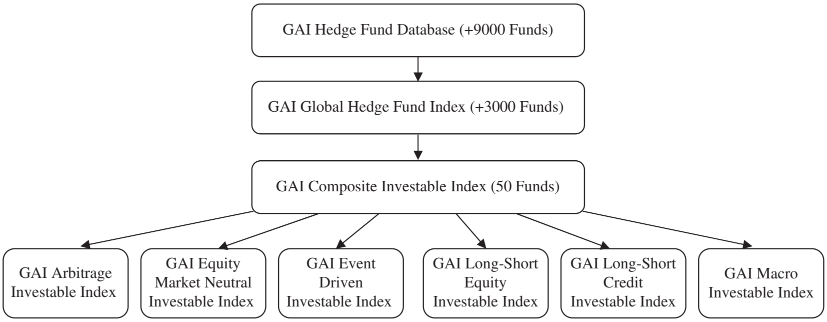


Figure 2.8 The family of GAI Investable Indices
Source: GAI

Indices follow those of the GAI Global Hedge Fund Indices in order to most accurately represent the current asset allocation within the hedge fund index. The GAI Investable Indices provide sophisticated investors with an investable benchmark for the hedge fund industry designed to track the various strategies of the GAI Global Hedge Fund Indices. Table 2.3 shows the family of GAI investable indices along with their respective replicating performance index.

The fund selection process begins with a quantitative review of the performance of all the funds in the GAI Global Hedge Fund Index. Each one is then ranked within their assigned strategy according to the Greenwich Value Score™ (GVS). The GVS is a multi-factor model designed to identify persistent top-quartile funds and provide a relative performance rating and risk assessment based on:

- Risk adjusted performance,
- Volatility,
- Downside risk, and
- Correlations.

Table 2.3 GAI Investable Indices and associated replicating index

GAI Investable Index Name	# Funds	GAI Replicating Index
Composite Investable Index (Mth)	50	Global Hedge Fund Index
Composite Investable Index (Qtr)	50	Global Hedge Fund Index
Arbitrage Investable Index	10	Global Arbitrage Index
Equity Market Neutral Investable Index	10	Global Equity Market Neutral Index
Event-Driven Investable Index	10	Global Event-Driven Index
Futures Investable Index	10	Global Futures Index
Long-Short Credit Investable Index	10	Global Long-Short Credit Index
Long-Short Equity Investable Index	20	Global Long-Short Equity Index
Macro Investable Index	10	Global Macro Index

Source: GAI

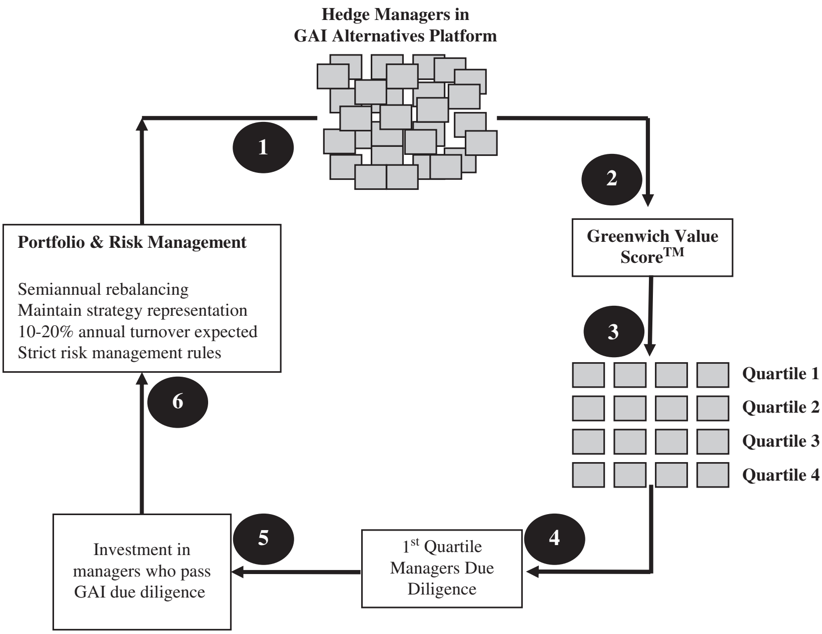


Figure 2.9 Schematic of the fund selection process
Source: GAI

Potential candidates are then examined on a qualitative basis and subjected to a rigorous due diligence procedure, including a detailed review of the hedge fund model to determine strategy, style and expected returns whilst also highlighting a manager's ability to effectively create alpha and consistently deliver positive returns. During this process, an in-depth analysis of the fund strategy and risk control procedures is also undertaken as well as stress testing at various points within the fund's performance (see Figure 2.9).

In addition to meeting the above quantitative and qualitative processes, the constituent fund must also have a minimum one-year track record and AuM exceeding $50m (overall company assets greater than $100m).

2.2.6 Morningstar Alternative Investment Center (http://www.morningstar.com/advisor/alternative-investments)

In September 2008, Morningstar acquired the Morgan Stanley Capital International (MSCI) Hedge Fund Index family and agreed to use their industry-leading categorisation and index construction methodology. The Morningstar MSCI Hedge Fund Indices currently consist of over 190 indices.

2.2.6.1 MSCI Hedge Fund Classification Standard

In July 2002, MSCI launched the MSCI Hedge Fund Classification Standard (HFCS), one of the most comprehensive hedge fund classification models to date. The HFCS uses three Primary Characteristics, namely the hedge fund investment style, asset class and geography in order to classify funds and define hedge fund strategies. In addition, Secondary Characteristics are defined through the Global Industry Classification Standard (GICS), and cover capitalisation size for equity-oriented strategies, and fixed income focus for credit-oriented strategies (see Table 2.4). The investment process determines the approach managers use to select investments in order to generate returns and manage risk. They are grouped into five broad categories, or investment process groups, i.e. Directional Trading, Relative Value, Security Selection, Specialist Credit and Multi-Process. The indices are equal weighted at all four levels of aggregation and asset weighted at the two highest levels as can be seen in Table 2.5.

Classification schemes attempt to group hedge funds based on their strategy and characteristics; however, such meaningful groupings are challenging due to the inherent heterogeneity of hedge fund investing. Such diversification does not allow the development of a simple group-based system and any classification is likely to be subject to a large degree of subjectivity. The MSCI HFCS attempts to overcome this problem by attempting to capture the multi-dimensional nature of hedge fund investing through the use Primary and Secondary Characteristics, as described above, to more accurately and effectively identify and define hedge fund strategies. The MSCI HFCS strives to offer a balance between a suitable level of detail which permits an accurate classification of a large number of strategies whilst also allowing an intuitive understanding of its interpretation and implementation.

2.2.6.2 MSCI Investable Indices

In addition to the family of MSCI Hedge Fund Indices, MSCI developed an index construction methodology to build a range of MSCI Investable Hedge Fund Indices. Their aim was to develop a set of replicable and tradable hedge fund indices reflecting the aggregate performance of a diversified range of hedge fund strategies. Such objectives require that the investable hedge fund indices be based on well-diversified managed account platforms (or *Platforms*) of hedge fund investments offering more frequent pricing and liquidity than would otherwise be possible.

An Investable Hedge Fund Reference Framework (HFRF) is designed to establish which hedge fund investment processes and strategies will be represented in the investable index. The HFRF is further supplemented by a series of index calculations, maintenance rules and guidelines relating to segment

Table 2.4 MSCI Primary and Secondary Characteristics breakdown

Primary Characteristics					Secondary Characteristics		
Investment Process			Geography				
Process Group	Process	Asset Class	Area	Region	GICS Sector	Fixed Income Sector	Market Cap
Directional Trading		Commodities	**Developed Markets**		Consumer Discr.	Asset-Backed	Small
	Discretionary	Convertibles		Europe	Consumer Staples	Gov. Sponsored	Small & Mid Cap
	Tactical	Currencies		Japan	Energy	High Yield	Mid & Large Cap
	Systematic	Equity		North America	Financial	Investment Grade	No Size Focus
	Multi-Process	Fixed Income		Pacific ex Japan	Health Care	Mortgage-Backed	
		Diversified		Diversified	Industrials	Sovereign	
Relative Value			**Emerging Markets**		IT	No Sector Focus	
	Arbitrage			EMEA	Materials		
	Merger Arb.			Asia Pacific	Telecom Services		
	Statistical Arb.			Latin America	Utilities		
	Multi-Process			Diversified	No Industry Focus		
Security Selection			**Global Markets**				
	Long Bias			Europe			
	No Bias			Asia ex Japan			
	Short Bias			Asia			
	Variable Bias			Diversified			
Specialist Credit							
	Credit Trading						
	Distressed						
	Private						
	Multi-Process						
Multi-Process Group							
	Event Driven						
	Multi-Process						

Source: MSCI

Table 2.5 The MSCI Hedge Fund Index classification structure

					Weighting
Hedge Fund Composite Indices					Asset & Equal
Process Group	**Process Group**	**Process Group**	**Process Group**	**Process Group**	Asset & Equal
Directional Trading	Relative Value	Security Selection	Specialist Credit	Multi-Process	
Investment Process	**Investment Process**	**Investment Process**	**Investment Process**	**Investment Process**	Equal
Discretionary Trading Tactical Allocation Systematic Trading Multi-Process	Arbitrage Merger Arbitrage Statistical Arbitrage Multi-Process	Long Bias No Bias Short Bias Variable Bias	Distressed Securities Long-Short Credit Private Placements Multi-Process	Event Driven Multi-Process	
Strategy Indices	**Strategy Indices**	**Strategy Indices**	**Strategy Indices**	**Strategy Indices**	Equal

Source: MSCI

diversification, fund eligibility, concentration, investment capacity and fund and segment weight allocation. The MSCI Hedge Fund Composite Index (HFCI) is used as a proxy for the hedge fund universe. The HFCI is an equally weighted index that measures the performance of a diverse portfolio of hedge funds (AuM greater than $15m) across the range of available investment strategies. Figure 2.10 shows a schematic of the MSCI investable index construction methodology.

A review of the investable hedge fund index is conducted on a quarterly basis where funds are added or deleted and adjustments made to the constituent weights for available investment capacity. Segment weights are also realigned to the HFRF and checked to ensure that they still adhere to the general index construction and maintenance principles.

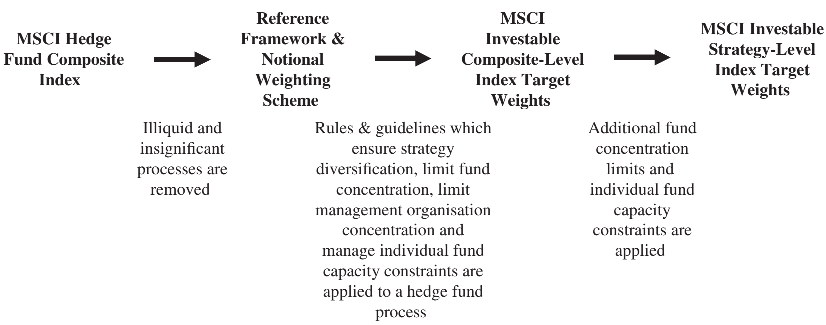


Figure 2.10 A schematic of the MSCI investable index construction methodology
Source: MSCI

2.2.7 EDHEC Risk and Asset Management Research Centre (www.edhec-risk.com)

As detailed above, the different hedge fund indices available in the market conform to a range of selection criteria and are developed through a variety of construction methodologies attached to a variety of commercial databases. With such inherent heterogeneity in the hedge fund industry, investors cannot rely on competing hedge fund indices to obtain a *true and fair* representation of hedge fund performance. As a response to this lack of representativeness and *purity* in the hedge fund industry, EDHEC Risk and Asset Management Research Centre (EDHEC) introduced a novel *index of indices* idea to alleviate the problem. Such *Alternative Indices* were first discussed in an EDHEC working paper by Amenc and Martellini (2003) who argued that due to the impossible nature of ascertaining the *best* index available on the market, a better construction method is to use a combination of competing indices to determine a more robust and representative industry benchmark. Since competing indices are based on different sets of hedge funds in their composition, the resulting portfolio of indices must be more exhaustive than any of the competing indices it is derived from.

EDHEC use a method of factor analysis to develop a set of hedge fund indices that are the best possible one-dimensional summaries of information within the competing indices for a given investment style. Such an analysis usually involves applying the method of Principal Component Analysis (PCA). Mathematically, this involves finding the first component of competing indices using their historical performance data. The first principal component has a built-in element of *optimality*,[12] because there is no other linear combination of competing indices that implies a lower information loss. Indeed, information is lost where the heterogeneity in the competing hedge funds' construction is the most severe. Since competing indices are affected in different ways by measurement biases, determining the linear combination of competing indices that implies a maximisation of the variance explained, leads implicitly to a minimisation of the bias.

2.3 DATABASE AND INDEX BIASES

Hedge funds report monthly returns to commercial databases on a voluntary basis; such participation means that only a portion of the hedge fund universe is observable and represented. For example, hedge fund managers will tend to report to databases only when their performance is good and may stop reporting once they become less attractive. This effect creates a variety of biases in the databases which can lead to vendors publishing misleading and incomplete return statistics, especially when considering index construction

[12] *Optimality* refers to stable, more representative, easy to replicate, non-commercial and with fewer biases.

and their subsequent publication. Such a problem does not occur in the mutual fund industry, where public disclosure of net asset values is enforceable by law, causing a natural convergence of the universe and database of mutual funds. However, it is well known that hedge fund performance data and their benchmarks inherit measurement biases from the databases on which they are based. As a consequence, it is particularly important to be completely aware of the origin and the consequence of potential measurement biases.

2.3.1 Survivorship Bias

If a database contains only information on funds that are active and report regularly to a database (i.e. *live* funds), then a *survivorship bias* can be introduced into calculated performance measures and indices constructed using such data. It is, however, important to distinguish between funds that have simply exited a database (i.e. *defunct* funds) and those that have ceased operation altogether due to bankruptcy or liquidation (i.e. *dead* funds). A defunct fund is a fund that was in a database but ceases to report information to the vendor for whatever reason (e.g. merger); a dead fund is one that is known to have terminated operations and closed down completely. Clearly, a dead fund must also be a defunct fund, but a defunct fund need not necessarily be a dead fund. Other funds that are defunct but not dead are those that have reached their capacity and no longer require additional capital or the need to attract new investors. Alternatively, the fund manager may believe their performance is so good that their investment style must remain private and no longer wish to provide sensitive information to a database that may be publicly available.

The effect of survivorship bias has been well known in the mutual fund industry for some time and is fairly straightforward to determine. The standard method of determining the survivorship bias, first proposed by Malkiel (1995), is to obtain the universe of all mutual funds that are active during a given time period. The average return of all funds is compared with that of the surviving funds at the end of the period. The return difference is survivorship bias. However, survivorship bias in hedge funds cannot be measured directly because the universe of hedge funds is not readily observable. Survivorship bias can only be estimated using a sample of hedge funds in a database. Technically, over any sample period, if a complete record of defunct funds is available, survivorship bias can be estimated through tedious data manipulation. The problem is in verifying the completeness of historical records on defunct hedge funds. The magnitude of the survivorship bias generally depends on two parameters, namely:

1. The attrition rate,[13] and
2. The average returns difference between surviving and dead funds.

[13] The *attrition rate* is the percentage of hedge funds that fail over a given time period, e.g. a year.

The hedge fund industry is all too aware that the exclusion of a fund from a database can lead to an upward bias in hedge fund returns and an understated historical risk. Xu, Liu and Loviscek (2009) studied a major commercial database of hedge funds from January 1994 to March 2009 and found that hedge fund returns were generally much worse than the industry would have us believe since many failing funds stop reporting their performance. The gap in returns between these failing funds and others averaged 0.54% a month, or about 6% per annum. This is an extremely high figure and suggests that average reported hedge fund returns set an unrealistic expectation for hedge fund investors.

As regards defunct funds, there exists another type of bias, namely liquidation bias. This is a result of the fund stopping reporting hedge fund performance to the database several months prior to the final liquidation value of the fund whilst they concentrate on winding down their operations. This generally causes an upward bias in the returns of defunct funds. The opposite of liquidation bias is participation bias. This bias can occur with a successful hedge fund manager who closes his fund and stops reporting his results because he no longer needs to attract new capital.

Another related concept to survivorship bias, and again a result of voluntary reporting, is self-selection bias. Funds not performing well can hide bad results in order to avoid investors withdrawing their money whereas funds performing well may wish to protect their investment strategies, and stop the inflow of capital by ceasing to report and closing the fund. In fact, some hedge fund managers choose to be included in a database, period by period, depending on the fund's performance.[14] The fact that hedge fund managers can choose when to participate in the database leads to a self-selection bias. Such a bias is almost impossible to determine let alone measure.

2.3.2 Instant History Bias

For many databases, there is often a sizeable lag between a fund's inception date and the date at which the fund returns are submitted to a particular database. This time lag typically corresponds to the hedge fund incubation period (12–18 months) where the performance of the fund is evaluated using *seed* investment before being publicly offered to investors in order to attract further capital. Once a manager is in a position to submit attractive performance data to a database, they naturally choose the start date that shows the hedge fund in the most positive light. The database vendor introduces an "instant history" bias into the data when they decide to *backfill* the data to show the historic performance of the hedge fund even though such data was not available when the database was established. Clearly, such funds are likely to be those that offer higher

[14] Many hedge fund index compilers prohibit this practice and insist that managers regularly and timely submit performance data in order to be considered for inclusion in an index.

returns and therefore backfilled data will invariably inflate the performance of the fund in the earlier days. Different databases handle the issue of *backfilling* data differently, and as a result, the impact of this bias varies between vendors. To reduce or avoid this bias some vendors do not backfill returns at all. Others, however, backfill only a few months. In any event, performance returns obtained from databases should be handled with care.

2.4 BENCHMARKING

The development of traditional indices rests on the assumptions that the underlying instruments are homogenous, and that an investor follows a simple *buy and hold* strategy. Traditional indices are constructed to represent the return of the *market portfolio*; an asset-weighted combination of all investable instruments in that asset class or a suitably equivalent proxy. These indices are designed to directly define the risk premium available to investors willing to expose themselves to the systematic risk of the asset class. For example, an investor buying the Dow Jones will be exposed to a broad range of market risks based on 30 US large cap equities, i.e. there exists a general equilibrium model. However, such a model is still absent from the hedge fund world. In the early years of hedge funds, investment committees established a type of hedge fund performance indicator based on the idea of an *absolute return*, loosely defined as a flat rate of return obtainable under any market condition (i.e. 14% p.a.) with no reference to a market average or *peer* group measure. As the market has become more challenging and competitive for hedge fund managers, benchmarking to an absolute return in its purest sense has become practically impossible. Hedge fund managers naturally focus their efforts on liquid markets, where trading opportunities and leverage are readily available. Thus, as the dynamics in the global markets change, the nature of hedge funds in the market also changes. Benchmarking such a dynamic industry in itself is an arduous task, and the difficulty is further exemplified by the fact that the hedge funds that constitute the benchmarks are drawn from a sample of funds managed by managers with diverse investment styles. For this reason, investors are increasingly relying on a range of hedge fund indices, as discussed above, along with their inadequacies, as their primary method of benchmarking. Benchmarking with an index only makes sense if the index has the following characteristics and attributes. It must be:

- Representative
- Rule-based
- Fully investable
- Transparent
- Diversified
- Timely reported, and
- Liquid.

Having a reliable benchmarking system is one of the biggest challenges that institutional investors face when selecting and evaluating hedge fund managers and their returns. Most institutional investors are interested in analysing how hedge fund strategies correlate with and compare to broad market indices for portfolio construction, optimisation and asset allocation purposes. As the hedge fund industry matures and becomes ever more driven by large institutional investment, benchmarking is sure to increase.

2.4.1 Tracking Error

The tracking error is a measure of how closely a portfolio follows the index to which it is benchmarked. The lower the tracking error, the more the fund resembles its benchmark's risk and return characteristics. In all cases, the benchmark is the measured position of neutrality for the hedge fund manager. If a manager were to simply follow the benchmark, the expectation would be that their performance should equal the performance of the benchmark, and their tracking error should be zero. The most common measure is the difference in the return earned by a portfolio and the return earned by the benchmark against which the portfolio is constructed. For example, if a particular hedge fund earns a return of 9.15% during a period when the particular benchmark produces a return of 9.07%, the tracking error is 0.08%, or eight basis points.

Tracking error may be calculated from historical performance data or estimated for future returns. The former is called *ex-post* and the latter *ex-ante* tracking error. Mathematically, tracking error (*TE*) can be defined in terms of the standard deviation (*SD*), such that:

$$TE = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(r_{N-1,N} - r_{N-1,N}^{bm}\right)^2} \tag{2.1}$$

where N is the total number of sample data points, $r_{N-1,N}$ is the hedge fund return and $r_{N-1,N}^{bm}$ is the benchmark return.

Many practitioners have argued that the quadratic form of the *TE* is difficult to interpret, and that hedge fund managers generally think in terms of linear and not quadratic deviations from a benchmark. In this case, *TE* in terms of mean absolute deviations (*MAD*) can be written as:

$$MAD = \frac{1}{N-1}\sum_{i=1}^{N}\left|r_{N-1,N} - r_{N-1,N}^{bm}\right| \tag{2.2}$$

In this chapter we have discussed the major hedge fund databases available in the market and the range of services and products they offer. We have highlighted the fact that the hedge fund industry is highly heterogeneous which makes the construction of a useful market index that comprises the available hedge fund universe very difficult. Moreover, the issues relating to the different ways in which hedge funds report to commercial databases creates a variety of data biases which can lead to vendors publishing misleading and incomplete performance measures. Non-investable hedge fund indices try to represent the performance of a sample of the hedge fund universe taken from a particular database, however, such vendors have diverse selection criteria and methods of index construction leading to many different published indices. More recently, many index providers have developed investable indices that offer a low-cost investment and exposure to the hedge fund industry. Nevertheless, one should be fully aware of the limitations and constraints before accepting a particular index has a valued industry benchmark.

In Chapters 1 and 2 we have covered the fundamentals and key issues related to the hedge fund industry. In the following chapters we will build and develop a quantitative and theoretical approach to modelling and analysing hedge funds. Chapter 3 begins this process by introducing the main statistical methods and techniques applied to hedge funds.

3 Statistical Analysis

In order for hedge fund managers to make informed decisions with regards to their investments it is essential that several key statistical analyses are performed. This will usually involve analysing a time series of periodic hedge fund returns to ascertain relevant statistical properties of the data in order to make critical inferences about the characteristics and performance of the hedge fund. Many visual and mathematical methods are available that allow hedge fund managers to understand the underlying data structure and identify potential anomalies that may need further investigation whilst also allowing managers to make better informed decisions. It is also important that a serious investor or hedge fund manager have a working knowledge of many of the probability and statistical concepts encountered in the industry so as to be confident and knowledgeable when explaining and discussing their hedge fund investment strategies to potential investors.

Chapter 3 covers the main concepts, principles and techniques employed in the statistical analysis of hedge fund returns. Both visual and theoretical methods are presented which show how to extract and interpret the informational content and underlying characteristics within a time series of hedge fund returns.

3.1 BASIC PERFORMANCE PLOTS

3.1.1 Value Added Index

Extracting any valuable information from a time series of hedge fund returns is practically impossible with a large set of data in tabular format. Consider the hypothetical monthly[1] returns for a CTA Index between 2008 and 2013, i.e. a total of 72 individual positive and negative returns.[2] A simple bar chart of the data, as shown in Figure 3.1, gives an instant visualisation of the historical performance of the CTA Index over the time period as well as identifying areas of highs and lows.

[1] Hedge funds can also report daily and weekly returns.

[2] A hedge fund may only provide a series of periodic Net Asset Values (NAV) in which case they can be converted into an equivalent series of periodic returns using $r_{t_1,t_2} = \frac{NAV_{t_2} - NAV_{t_1}}{NAV_{t_1}}$ where r_{t_1,t_2} is the return for the hedge fund between time t_1 and t_2 $(t_2 > t_1)$.

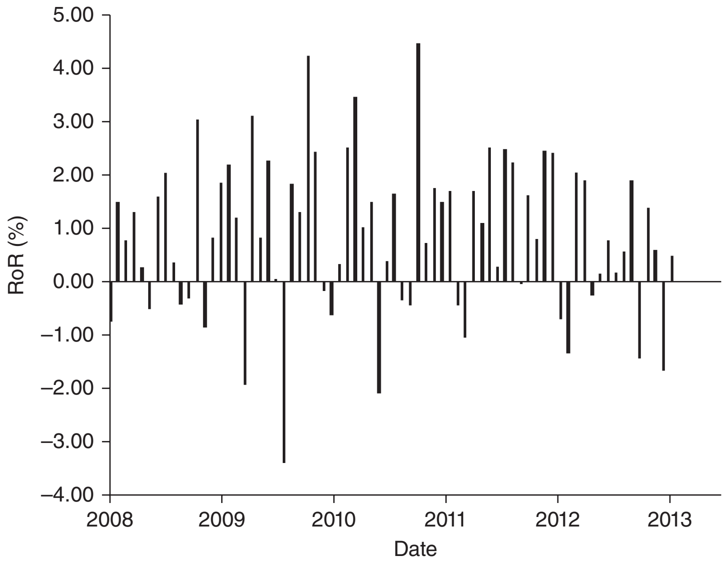


Figure 3.1 Bar chart of the monthly returns for a hypothetical CTA Index (2008–2013)

```
%File: barchart.m
%import XL data
[~,dates,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

n = size(returns,1); %# data points

figure; %create figure
sDate = datenum(dates(1)); %set start date for x-axis
eDate = datenum(dates(end)); %set end date for x-axis
xData = linspace(sDate,eDate,n);

bar(xData, returns); %plot bar chart
xlabel('Date'); %add x label
ylabel('RoR (%)'); %add y label
box off;
ytick = get(gca,'YTick'); %format axes
set(gca,'YTickLabel', sprintf('%.2f|',ytick))
datetick('x','yyyy','keeplimits');
```

Source 3.1 MATLAB® code for the bar chart in Figure 3.1

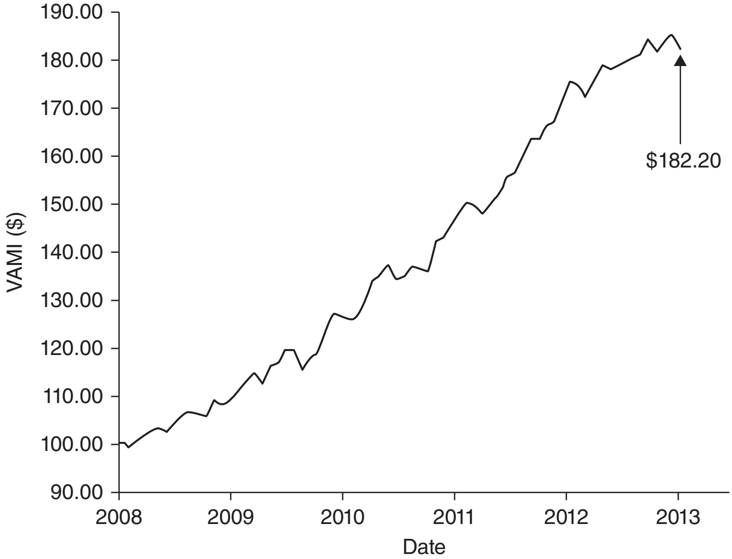


Figure 3.2 The VAMI plot for a hypothetical CTA Index of monthly returns (2008–2013)

The Value Added Index (VAI) is an index that shows the performance of a hypothetical $100[3] investment based on a set of returns for a hedge fund with a specific reporting frequency, e.g. daily, weekly, monthly. The can be calculated for daily, weekly or monthly returns depending on the reporting frequency. The VAI is given by:

$$\mathrm{VAI}_{t+1} = \mathrm{VAI}_t(1 + r_t) \tag{3.1}$$

The VAMI plot for the hypothetical CTA Index of monthly returns between 2008 and 2013 is shown in Figure 3.2. Source 3.2 shows the MATLAB® code of a user-defined function for calculating VAI.

Reading off the final value of the $100 investment in the hypothetical CTA Index over the five-year period gives a VAMI value of $182.20; however, it does not really tell us much about the performance of the CTA Index since 2008.

3.1.2 Histograms

A histogram is a graphical summary of the frequency distribution of a set of *empirical data*,[4] for example a time series of monthly hedge fund returns

[3] Some hedge fund managers prefer to use $1000 for the initial value.

[4] *Empirical data* are those gained by means of a series of observations or by experiment. Empirical measures are those determined from observed values, as opposed to those calculated using theoretical models.

```
%File: fVAI.m
%VALUE ADDED INDEX (VAI)
function vaidx = fVAI(x)
%---------------------------------------------------------------------
%x:       returns vector
%vaidx:   Value Added Index (VAI)
%---------------------------------------------------------------------

    n = size(x,1); %# data points
    fVAI = ones(1,n); %pre-allocate memory
    fVAI(1) = 100; %initialise to $100
    for i=1:n-1
        fVAI(i+1) = fVAI(i) * (1+x(i)./100); %calculate VAI
    end
    vaidx = fVAI;
end
```

Source 3.2 User-defined MATLAB® function for the calculation of VAI

```
%File: vami.m
%import XL data
[~,dates,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

n = size(returns,1); %# data points
y = fVAI(returns); %get VAMI values

figure; %create figure
sDate = datenum(dates(1)); %set start date for x-axis
eDate = datenum(dates(end)); %set end date for x-axis
xData = linspace(sDate,eDate,n);

plot(xData,y); %plot line chart
xlabel('Date'); %add x label
ylabel('VAMI ($)'); add y label
box off;
ytick = get(gca,'YTick'); %format axes
set(gca,'YTickLabel', sprintf('%.2f|',ytick))
datetick('x','yyyy','keeplimits');
annotation('textarrow',[.845,.845],[.7,.85],'String','$182.20'); %add
arrow & text
```

Source 3.3 MATLAB® code for the VAMI plot in Figure 3.2

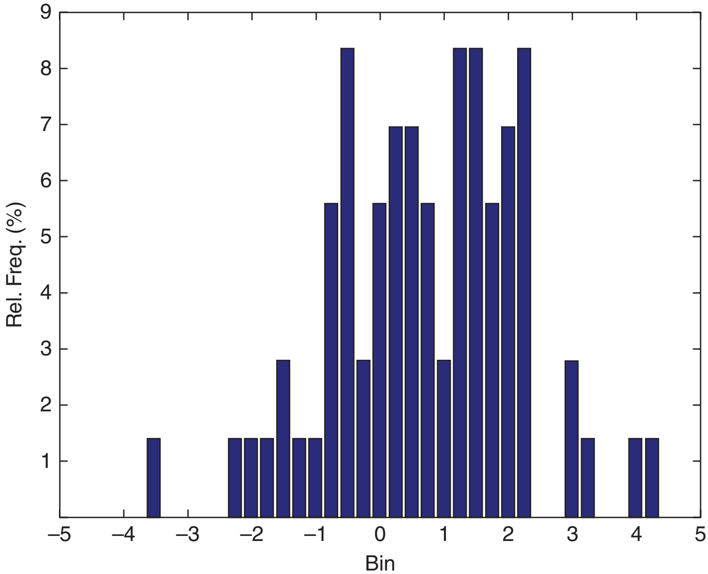


Figure 3.3 Histogram plot for a hypothetical CTA Index of monthly returns (2008–2013)

(or NAVs). The frequency is an absolute value in which each value represents the actual count of the number of occurrences of a particular value in a group of data. Relative frequencies involve *normalising* the frequencies through division of the absolute frequency by the total number of observations in the group of data. Knowing the maximum and minimum values of the monthly returns it is possible to set a range that encompasses all the monthly returns in the data set. This range is then divided into equal intervals (or *bins*) and an absolute frequency determined for each number of values that fall within each bin using the built-in MATLAB® function `histc()`. Once the absolute values have been determined, a relative frequency can be calculated by dividing each absolute frequency by the sum of absolute frequencies, i.e. 72 and a subsequent histogram of the frequency distribution plotted as in Figure 3.3. The MATLAB® code that implements the above methodology is shown in Source 3.4.

The histogram is a very useful representation of a set of hedge fund returns over a specific time period and gives a good visual representation of the shape of the distribution and highlights areas of high negative and positive returns.

3.2 PROBABILITY DISTRIBUTIONS

In order to understand probability distributions, it is important to understand what is meant by a *random variable*. When a numerical value of a variable

```
%File: histogram.m
%import XL file
[~,dates,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

mn = abs(min(returns)); %get min percent absolute
mx = abs(max(returns)); %get max percent absolute

figure; %create figure
y=returns; %set returns to y values

x=-4.5:0.25:4.5; %bins -4.5 to 4.5 intervals 0.25
freq = histc(y,x); %count frequency per bin
rfreq = (freq/sum(freq))*100; %relative frequency

bar(x,rfreq); %plot bar chart
xlabel('Bin'); %add x label
ylabel('Rel. Freq. (%)'); %add y label
```

Source 3.4 MATLAB® code for the histogram plot in Figure 3.3

is determined by an unknown (or chance) event, that variable is said to be random. Random variables can either be *discrete* or *continuous*. For example, suppose an experiment consists of flipping a coin six times and recording the number of heads that appear after each toss. The number of heads results from a *random process* (i.e. flipping the coin) and the actual number recorded is a value between 0 and 6, i.e. a finite integer value. Therefore, the number of heads is said to be a *discrete* random variable. Now suppose the same experiment is performed, but the average number of heads after flipping the coin six times is recorded. The *average* number of heads again results from a random process, however the actual number recorded can now be any value between 0 and 1, i.e. an infinite number of values. In this case, the average number of heads after six coin flips is said to be a *continuous* random variable.

A *probability distribution* describes all the possible values that a random variable can take within a given range. Probability distributions can also be discrete or continuous. For example, consider an experiment in which a coin is flipped two times and let a random variable[5] X represent the number of heads that occur. The four possible outcomes to this experiment are HH, HT, TH and TT so the discrete random variable X can only have values 0, 1 or 2. That is, the experiment can be described by a discrete probability distribution as shown in Table 3.1.

[5] Generally a random variable is denoted by an uppercase letter, e.g. X and the possible values of the random variable denoted by lowercase letters, e.g. $\{x_1, x_2, x_3\}$.

Table 3.1 The discrete probability distribution for the random variable X

Number of Heads	Probability
0	$\frac{1}{4}$
1	$\frac{1}{2}$
2	$\frac{1}{4}$
Total Probability	1

A continuous probability distribution differs from a discrete probability distribution since the probability that a continuous random variable will equal a certain value is always zero. That is, the continuous random variable can take on an infinite number of values. As a result, continuous probability distributions cannot be expressed in tabular format (as in Table 3.1) but have to be described in terms of a mathematical function known as a probability density function. However, before looking at such density functions, it is worthwhile taking the time to understand the difference between population and samples when discussing probability distributions and associated statistical measures.

3.2.1 Populations and Samples

A population includes every element from a set of possible observations (i.e. the entire data set), whereas a sample consists only of those elements drawn from the population. Depending on the sampling method, it is possible to derive any number of samples from a population (see Figure 3.4). Furthermore, a statistical measure associated with a population, such as mean or standard deviation, is known as a *parameter*; but a statistical measure associated with a particular sample of the population is called a *statistic*. However, most statistical measures in the finance world, although determined from samples of data, are often referred to as parameters.

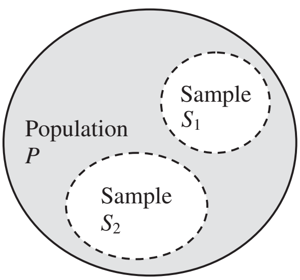


Figure 3.4 Several samples taken from the population of the data set

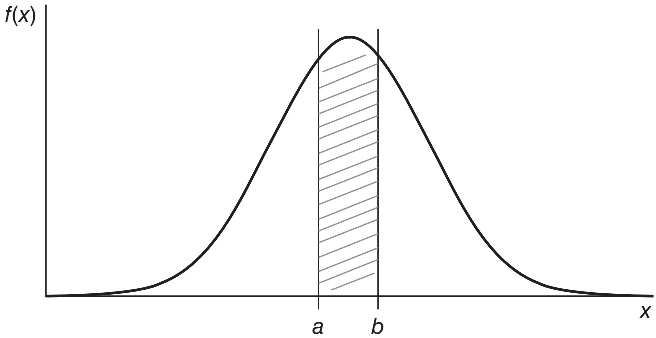


Figure 3.5 The PDF and interval $[a, b]$

A sampling method is the process of selecting a sample from the population. When considering random sampling, several properties must hold, for example:

- The population consists of N elements,
- The random sample consists of n elements, and
- All random samples of n elements are equally likely to occur.

Adhering to the above properties ensures that the chosen random sample is representative of the total population and that any statistical assumptions made about the random sample will be valid.[6]

3.3 PROBABILITY DENSITY FUNCTION

Given a continuous random variable, X, the PDF[7] of X is a function $f(x)$ such that for two numbers, a and b with $a \leq b$ we have:

$$P(a \leq X \leq b) = \int_a^b f(x)dx \tag{3.2}$$

Where $f(x) \geq 0$ for all x. So, the PDF of a continuous random variable is a function which when integrated over the limits a to b, gives the probability that the random variable will have a value within that given interval (or domain). More formally, the probability that X is a value within the interval $[a, b]$ equals the area under the PDF from a to b (see Figure 3.5).

[6] Although the sample may be subject to a *bias* compared with the population.
[7] The Probability Density Function (PDF) is also known as the probability distribution function, the probability mass function or density function.

As with discrete probability distributions, the total probability for a continuous probability distribution must also be one. Moreover, the total area under the PDF is always equal to one, that is:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{3.3}$$

3.4 CUMULATIVE DISTRIBUTION FUNCTION

The Cumulative Distribution Function (CDF), describes the probability that a random variable X with a given probability distribution will be found at a value less than or equal to x. The CDF is a function $F(x)$ of a random variable, X, for a number x, such that:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(s)ds \tag{3.4}$$

That is, for a given value x, $F(x)$ is the probability that the observed value of X will be at most x. Figure 3.6 shows the relationship between a typical PDF and CDF.

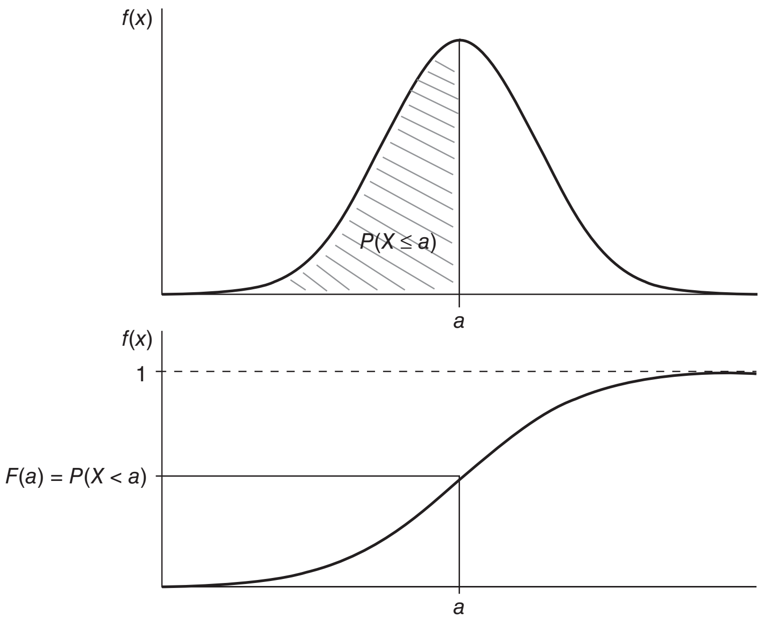


Figure 3.6 The relationship between the PDF and CDF

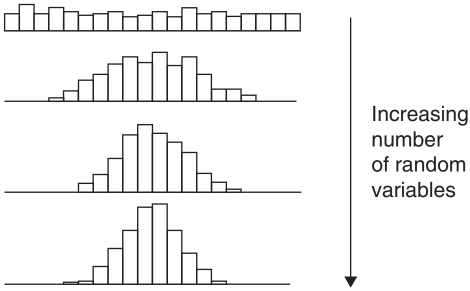


Figure 3.7 The central limit theorem

3.5 THE NORMAL DISTRIBUTION

A very popular probability distribution is the normal (or Gaussian[8]) distribution which has the following PDF:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{3.5}$$

where μ and σ are the mean and standard deviation of the probability distribution.

The normal distribution is considered the most prominent in probability and statistical theory (and in finance), and, in many real-life studies, probability distribution tends towards the normal distribution provided there are a sufficient number of random variables. Indeed, the central limit theorem (also known as the law of large numbers) states that the sum of a large number of independent and identically distributed (*iid*) random variables have an approximate normal distribution. The approximation improves as the number of random variables increases as illustrated in Figure 3.7.

In fact, the normal distribution is often used as a first approximation to a random variable that tends to *cluster* around a single mean value, i.e. μ. The graph of the normal distribution is symmetric about the mean and usually referred to as the *bell*-shaped curve as shown in Figures 3.7 and 3.8.

A normal distribution can be fully described by only two statistical parameters, the mean and standard deviation. When a random variable X is distributed with a mean μ and standard deviation σ, the normal distribution can be denoted in the following compact form:

$$X \sim N\left(\mu, \sigma^2\right) \tag{3.6}$$

[8] The Gaussian distribution is named after Carl Friedrich Gauss (1777–1855), a German mathematician and scientist who made major contributions in the fields of number theory, statistics, differential geometry, astronomy and optics.

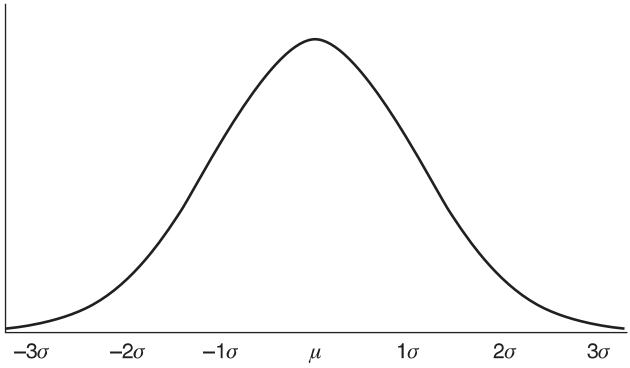


Figure 3.8 The normal distribution

The normal distribution is assumed to approximate the model for many financial time series, including the distribution of monthly hedge fund returns, although these can often deviate from normality and this must be taken into account when making inferences from such data.

3.5.1 Standard Normal Distribution

If $\mu = 0$ and $\sigma = 1$ the distribution is said to be a standard normal distribution (or z-distribution), such that:

$$X \sim N(0, 1) \tag{3.7}$$

With the following PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{3.8}$$

The normal random variable of a standard normal distribution is known as a z-score (or standard score). Every normal random variable X can be converted into a z-score using the following transformation:

$$z = \frac{X - \mu}{\sigma} \tag{3.9}$$

Where X, μ and σ are the random variable, mean and standard deviation from the original normal distribution, respectively. The z-score indicates the number of standard deviations above or below the mean, e.g. if the z-score is 2, then the original random variable X is 2 standard deviations above the mean. A negative z-score means that X is below the standard deviation by a certain amount.

3.6 VISUAL TESTS FOR NORMALITY

3.6.1 Inspection

A quick and simple visual test to determine how much a series of hedge fund returns deviate from normality is to plot the histogram of the empirical distribution against a fitted normal distribution. In this case, the bins are additionally used to calculate each normal value across the bin range using the built-in MATLAB® function `normpdf()`. So that the empirical values can be plotted against the normal values, it is necessary to *normalise* each value with respect to the total so that each complete set forms a unit area. Figure 3.9 is the plot of the empirical distribution against a normal distribution. The MATLAB® code that implements the above methodology is shown in Source 3.5.

Figure 3.9 shows that the distribution of monthly returns for the CTA Index has a reasonable bell-shaped curve which approximates the normal distribution. However, there seem to be several spikes that fall outside the normal which would need to be investigated further by considering higher moments of the monthly returns distribution, i.e. the skew and kurtosis measures.

3.6.2 Normal Probability Plot

Another very useful data visualisation technique and test for normality is the normal probability plot (or normal Q-Q plot). The built-in MATLAB® function

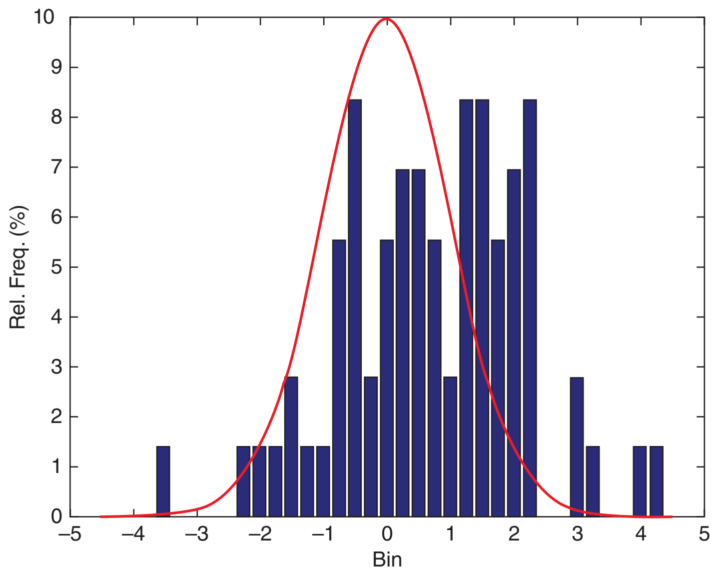


Figure 3.9 Empirical vs. normal distribution for a hypothetical CTA Index (2008–2013)

```
%File: combined.m
%import XL data
[~,dates,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

figure; %create figure
y=returns; %set returns to y values

x=-4.5:0.25:4.5; %bins -4.5 to 4.5 intervals 0.25
freq = histc(y,x); %count frequency per bin
rfreq = (freq/sum(freq))*100; %relative frequency
nfreq=normpdf(x); %standard normal frequency per bin
rnfreq = (nfreq/sum(nfreq))*100; %relative normal frequency

bar(x,rfreq); %plot chart combination
hold on;
plot(x,rnfreq,'r');
hold off;
xlabel('Bin'); %add x label
ylabel('Rel. Freq. (%)'); %add y label
```

Source 3.5 MATLAB® code for the plot in Figure 3.9

`normplot()` gives the calculation of the normal probability plot for a series of hedge fund returns as shown in Figure 3.10. The MATLAB® code that implements the normal probability plot is shown in Source 3.6. If the data are normal, the plot should fall more or less on a straight line between the data points. The hypothetical CTA Index data shows a very good fit to normality except around the tails of the distribution which can be seen from the departure of the data around the top and bottom of the straight line. Normal probability plots are often *S*-shaped indicating that the sample data is skewed or has heavier tails in relation to the normal distribution. A statistical measure of the *goodness-of-fit* is the correlation between the ordered data and *z*-scores. If the data are approximately normally distributed then the correlation should be a high positive value. Normal probability plots are easy to construct and interpret with the added advantage that outliers within the data are easily identified.

```
%File: normalplot.m
%import XL file
[~,dates,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

figure; %create figure
normplot(returns); %plot normal probability
```

Source 3.6 MATLAB® code for the normal probability plot in Figure 3.10

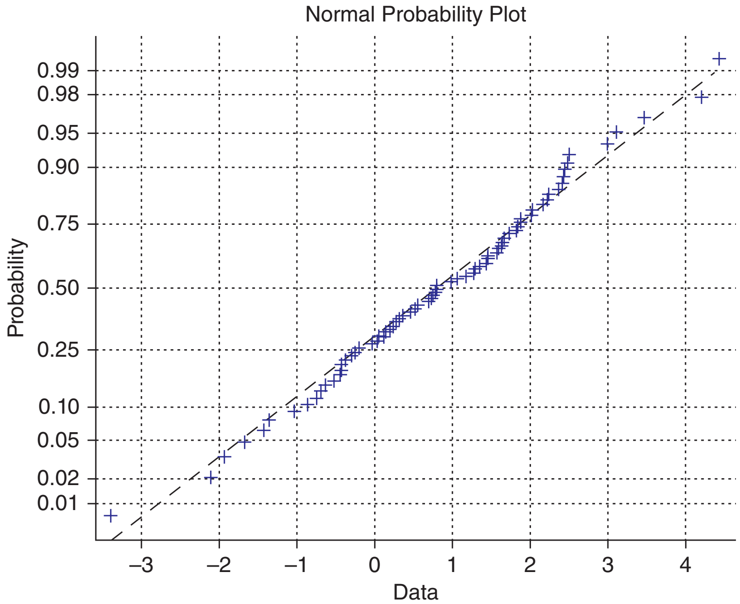


Figure 3.10 Normal probability plot for a hypothetical CTA Index (2008–2013)

3.7 MOMENTS OF A DISTRIBUTION

3.7.1 Mean and Standard Deviation

The mean ($m1$) and standard deviation ($m2$) are the first and second moments of a probability distribution and without doubt the two most quoted statistical measures used in finance.[9] The mean measures the average value of the distribution of a random variable and the standard deviation is the dispersion (or spread) of these values around the mean. The dispersion of returns around the mean is generally considered the amount of risk associated with a hedge fund. That is, the larger the standard deviation, the greater the potential hedge fund risk. For this reason, market practitioners often refer to the standard deviation as the volatility of the hedge fund.

For both the mean and standard deviation it is important to distinguish between individual population and sample measures. The mean for the population is denoted by μ and for the sample by $\bar{x}$. Similarly, the number of independent observations in a population is defined by N and for a sample

[9] Such statistical measures are generally known as *point estimates* since they use a sample data set to determine a single value (or statistic) which is a *best guess* for an unknown population parameter. Point estimates can be contrasted with *interval estimates*, e.g. confidence intervals.

```
%File: fMean.m
%MEAN
function m1 = fMean(x,f)
%------------------------------------------------------------------
%x:      returns
%f:      reporting frequency
%m1:     mean
%------------------------------------------------------------------

    m1 = mean(x).*f;
end
```

Source 3.7 User-defined MATLAB® function for the calculation of mean

taken from the population by n (where $n < N$). Mathematically, the mean for a population and sample are given by:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \tag{3.10}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3.11}$$

To get the annualised figure, the original mean is multiplied by the frequency representing the original time period, e.g. 12 for monthly and 4 for quarterly[10] as shown in Source 3.7.

The standard deviation for the population is denoted by σ and for the sample by s. Mathematically, the standard deviation[11] for a population and sample are given by:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} \left(x_i - \mu\right)^2}{N}} \tag{3.12}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n-1}} \tag{3.13}$$

[10] When considering daily returns, practitioners generally assume there are 252 trading days in a year.

[11] The standard deviation squared is known as the *variance*, however, the resulting formula gives a value in terms of squared units, e.g. returns squared ($\%^2$). Taking the square root of the variance gives the standard deviation and the correct units for returns (i.e. %).

```
%File: fStd.m
%STANDARD DEVIATION
function m2 = fStd(x,flag,f)
%---------------------------------------------------------------------
%x:      returns
%flag:   0 = sample, 1 = population
%f:      reporting frequency
%m2:     standard deviation (sample or population)
%---------------------------------------------------------------------

    m2 = std(x,flag).*sqrt(f);
end
```

Source 3.8 User-defined MATLAB® function for the calculation of standard deviation

Box 3.1 Square Root Rule – Standard Deviation

If a series of hedge fund returns are quoted in monthly or quarterly figures then they can be transformed into an equivalent annualised series using the so-called *square root rule*. To get the annualised figure, the original standard deviation is multiplied by the square root of the frequency representing the original time period, e.g. 12 for monthly and 4 for quarterly as shown in Source 3.8. More formally:

$$\sigma_{\text{annual}} = \sigma_{\text{monthly}} \times \sqrt{12}$$
$$\sigma_{\text{annual}} = \sigma_{\text{quarterly}} \times \sqrt{4}$$

One of the major advantages of using standard deviation, apart from the ease of calculation, is that it gives a direct measure of the riskiness of the distribution of returns for a hedge fund. However, most hedge fund returns do not exhibit a normal distribution, i.e. the familiar bell-shaped curve, but can be skewed or stretched in some way. In order to fully define a probability distribution it is necessary to investigate higher moments of the distribution, such as skewness and kurtosis.

3.7.2 Skew

Skew ($m3$) is the third moment of a probability distribution and measures the degree of asymmetry or skew of a distribution. Skew can be either zero, positive or negative. Positive (or right) skew indicates a distribution with an asymmetric tail extending toward more positive values, whereas with negative (or left) skew the tail extends toward more negative values (see Figure 3.11). The skew of a normal distribution is zero. Hedge fund returns are generally assumed to have

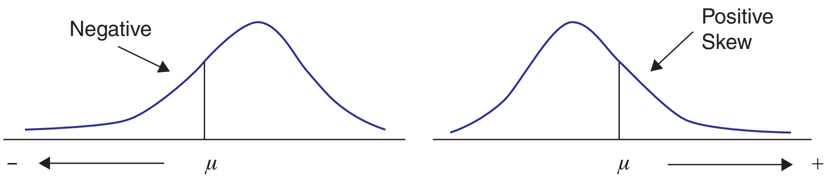


Figure 3.11 Positive and negative skew

either positive or negative skew; knowing the direction of the skew can help fund managers estimate whether a given (or future) return or price will be larger or smaller than the mean value. That is, if a particular return distribution is skewed, which is generally the case; there is a greater probability that the returns will be either higher or lower than that of a normal distribution. Unlike mean and standard deviation, skew has no units but is a pure number, like a *z*-score.

Mathematically, the sample skew, s is given by:

$$s = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s} \right)^3 \tag{3.14}$$

Box 3.2 Square Root Rule – Skew

To get the annualised figure, the original skew is divided by the square root of the frequency representing the original time period, e.g. 12 for monthly and 4 for quarterly as shown in Source 3.9. More formally:

$$s_{\text{annual}} = s_{\text{monthly}} / \sqrt{12}$$
$$s_{\text{annual}} = s_{\text{quarterly}} / \sqrt{4}$$

```
%File: fSkew.m
%SKEW
function m3 = fSkew(x,flag,f)
%----------------------------------------------------------------------
%x:      returns
%flag:   0 = sample, 1 = population
%f:      reporting frequency
%m3:     skew (sample or population)
%----------------------------------------------------------------------

    m3 = skewness(x,flag)./sqrt(f);
end
```

Source 3.9 User-defined MATLAB® function for the calculation of skew

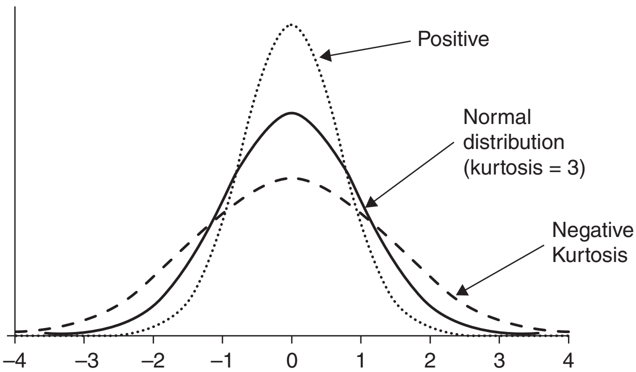


Figure 3.12 Positive, negative and zero kurtosis

3.7.3 Kurtosis

Kurtosis ($m4$), the fourth moment of a probability distribution, measures the degree of peakedness or flatness of a distribution compared to the normal distribution. When calculating kurtosis it is generally assumed that excess kurtosis is being considered. Positive kurtosis (or leptokurtic) indicates a relatively peaked distribution and heavy tails with more extreme values whilst negative kurtosis (platykurtic) refers to a flatter distribution with thinner tails and relatively fewer extreme values. A distribution with a kurtosis of three is known as a mesokurtic distribution,[12] e.g. the normal distribution (see Figure 3.12). A higher kurtosis usually indicates that the variability in the data is due to a few extreme variations from the mean, rather than many relatively small differences. As with skew, kurtosis also has no units.

Mathematically, the sample excess kurtosis, k is given by:

$$k = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)} \tag{3.15}$$

To get the annualised figure, the original kurtosis is divided by the frequency representing the original time period, e.g. 12 for monthly and 4 for quarterly as shown in Source 3.10.

Source 3.11 shows the MATLAB® code and results for the calculation of the four moments of the distribution of monthly returns for the hypothetical CTA Index (2008–2013). Note that both the skew and kurtosis have been adjusted for bias, i.e. using the built-in MATLAB® functions `skewness(x,0)` and `kurtosis(x,0)`.

[12] Rather than saying the kurtosis = 3 for a normal distribution, some talk instead of the *excess kurtosis* being zero, i.e. excess kurtosis = kurtosis - 3.

```
%File: fKurt.m
%EXCESS KURTOSIS
function m4 = fKurt(x,flag,f)
%-----------------------------------------------------------------------
%x:      returns
%flag:   0 = sample, 1 = population
%f:      reporting frequency
%m4:     excess kurtosis (sample or population)
%-----------------------------------------------------------------------

    m4 = (kurtosis(x,flag)-3)./f;
end
```

Source 3.10 User-defined MATLAB® function for the calculation of kurtosis

```
%File: moments.m
%import XL data
[names,~,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

%if monthly returns -> reporting frequency f = 12 for annual moments
f = 12;

m1 = fMean(returns,f); %annual mean (%)
m2 = fStd(returns,0,f); %annual standard deviation (%)
m3 = fSkew(returns,0,f); %annual skew (s)
m4 = fKurt(returns,0,f); %annual excess kurtosis (k)

OUTPUT:
```

	Mean (%) (m1)	Std. Dev. (%) (m2)	Skew (m3)	Kurtosis (m4)
CTA Index	10.245	5.111	-0.050	0.024

Source 3.11 MATLAB® code for the calculation of the four moments of the distribution of returns for the hypothetical CTA Index (2008–2013)

The annual mean and standard deviation (or volatility) for the hypothetical CTA Index are 10.245% and 5.111%, respectively. It is also evident that the hypothetical CTA Index has a degree of negative skew (–0.050) and positive kurtosis (0.024) which is a common characteristic of hedge fund return distributions.

3.8 COVARIANCE AND CORRELATION

Both covariance and correlation are related measures that indicate the degree of variation between two sets of random variables, for example a set of hedge fund returns and the market benchmark, e.g. S&P 500 Index. More formally, given any pair of random variables, X_i and X_j, their covariance is denoted by cov(X_i, X_j) or, in matrix form, by Σ_{ij}. By definition, the covariance is a symmetric matrix,[13] i.e. $\Sigma_{ij} = \Sigma_{,ji}$. Also, the covariance of any element X_i with itself is the variance, that is:

$$\mathrm{cov}(X_i, X_i) = \mathrm{var}(X_i) \tag{3.16}$$

The covariance matrix, Σ can be written as:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & . & . & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & . & . & . \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \Sigma_{n1} & . & . & . & . & \Sigma_{nn} \end{pmatrix} \tag{3.17}$$

To obtain a more direct indication of how two random variables co-vary, a correlation measure can be used. The correlation is simply a scaled version of the covariance with relation to the standard deviations of the two sets of random variables. There are several measures for the correlation, often denoted ρ (for a population) or r (for a sample), that indicate the degree of variation between two random sets of variables. The most common of these is the Pearson[14] product-moment correlation coefficient, which is relevant only to linear relationships between two sets of random variables. The correlation is $+1.0$ in the case of a perfectly positive (increasing) linear relationship, -1.0 for a perfectly negative linear relationship and a value between -1.0 and 1.0 in all other cases (see Figure 3.13). Note that if the elements of X and Y are independent then the correlation is zero, however, this does not indicate a lack of variation between X and Y, only that there is no linear relationship between them. However, there may be some other form of relationship between X and Y, such as a curvilinear[15] one.

Given any pair of random variables, X_i and X_j, the correlation ρ_{ij} is defined as:

$$\rho_{ij} = \frac{\mathrm{cov}\left(X_i, X_j\right)}{\sigma_i \sigma_j} \tag{3.18}$$

[13] The covariance matrix must also be *positive definite*, i.e. a matrix that is analogous to a positive real number.

[14] Karl Pearson (1857–1936) was an extremely influential English mathematician who has been most cited for the establishment of the field of mathematical statistics.

[15] A curvilinear relationship indicates that the relationship between the two sets of random variables may be curved.

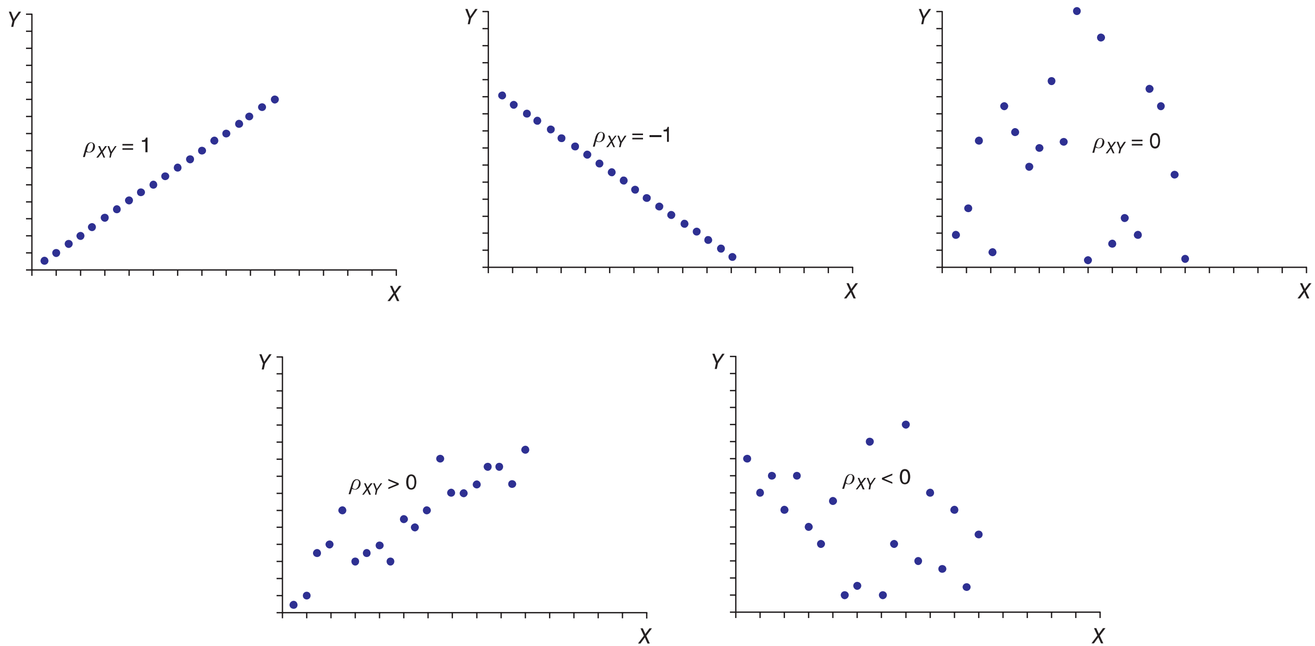


Figure 3.13 Some typical correlation plots

Where σ_i and σ_j are the standard deviations of X_i and X_j, respectively. The correlation is defined only if σ_i and σ_j are finite and nonzero. As with covariance, ρ_{ij} is more often seen in the form of a correlation matrix[16] ρ, that is:

$$\rho = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdot & \cdot & \rho_{1n} \\ \rho_{21} & 1 & \rho_{23} & \cdot & \cdot & \cdot \\ \rho_{31} & \rho_{32} & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{n1} & \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix} \tag{3.19}$$

Clearly, correlation inherits the symmetric property of covariance, i.e. $\rho_{ij} = \rho_{,ji}$. Note that $\rho_{ii} = 1.0$ for $i = j$. The correlation can also be transformed into the following computational form:

$$\rho_{XY} = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2}} \tag{3.20}$$

Where $\bar{x}$ and $\bar{y}$ are the mean values of X and Y.

Since correlation measures the relative strength of variability between two sets of random numbers it is a very useful tool for determining the degree of diversification within a portfolio of hedge funds, i.e. a fund of hedge funds (FoHF). Moreover, the correlation matrix representing a portfolio of hedge funds should ideally have relatively low values indicating a well-diversified portfolio of funds. Source 3.12 shows the individual correlations using the built-in MATLAB® function `corr()` for 10 hypothetical hedge funds of monthly returns. The correlation values range between –0.303 and 0.266 with the majority closer to zero indicating a small degree of positive and negative variability between these sets of hedge funds. Obviously, the correlation of a hedge fund with itself is one.

Investing in hedge funds and hedge fund strategies that have low correlations with each other is an ideal way for a fund manager to maximise the potential returns of an investment under a wide range of economic and market conditions.[17] Clearly, creating a well-diversified portfolio of low correlated hedge funds is an extremely valuable process for constructing profitable FoHFs investment strategies.

[16] As with covariance matrices, correlation matrices must also be positive definite.

[17] Assuming that correlations remain stable over time, which is not always the case especially in times of market turmoil and stress.

```
%File: correlation.m
%import XL data
[names,~,returns] = getXLData('hfma_matlab_data.xlsx','10 Hedge Funds');

%calculate correlation matrix
correl = corr(returns);

OUTPUT:
```

	CTA1	CTA2	CTA3	GM1	GM2	LS1	LS2	LS3	MN1	MN2
CTA1	1.000	-0.072	0.017	-0.008	-0.164	-0.045	0.044	0.137	0.109	0.266
CTA2	-0.072	1.000	-0.246	0.053	0.026	-0.069	0.102	0.006	-0.074	-0.303
CTA3	0.017	-0.246	1.000	0.043	-0.005	-0.059	-0.099	0.058	0.258	0.088
GM1	-0.008	0.053	0.043	1.000	0.011	0.044	-0.108	-0.082	0.201	-0.065
GM2	-0.164	0.026	-0.005	0.011	1.000	-0.072	0.020	-0.131	-0.013	-0.233
LS1	-0.045	-0.069	-0.059	0.044	-0.072	1.000	0.193	-0.091	0.033	-0.012
LS2	0.044	0.102	-0.099	-0.108	0.020	0.193	1.000	0.093	-0.090	0.084
LS3	0.137	0.006	0.058	-0.082	-0.131	-0.091	0.093	1.000	0.135	0.147
MN1	0.109	-0.074	0.258	0.201	-0.013	0.033	-0.090	0.135	1.000	-0.050
MN2	0.266	-0.303	0.088	-0.065	-0.233	-0.012	0.084	0.147	-0.050	1.000

Source 3.12 MATLAB® code for calculating the correlation matrix for the 10 hypothetical hedge funds (2008–2013)

3.9 LINEAR REGRESSION

Since the majority of hedge fund strategies involve investments in underlying financial instruments, a hedge fund manager is heavily exposed to the risks involved in using such instruments, or to so-called *market factors*. Moreover, many hedge fund managers invest in a whole range of different financial instruments in order to effectively implement a particular strategy and incorporate a degree of diversification into their portfolios. In order to help with this process, fund managers can identify the risk exposure of a given strategy by employing a combination of correlation and regression techniques. That is, for a certain hedge fund strategy or index, a set of correlated market factors can be identified and the strength of their relationship further quantified through regression analysis. In this way, a relationship can be ascertained between a dependent variable (e.g. hedge fund strategy or index) and a set of independent variables (i.e. correlated market factors). Since many fund managers rely on the use of derivative instruments and varying degrees of leverage,[18] correlation and regression analysis can only explain part of the risk exposures faced by such hedge funds. In these cases, a thorough knowledge of the particular strategy and market

[18] The use of derivatives and leverage introduces an element of *non-linearity* into the analysis.

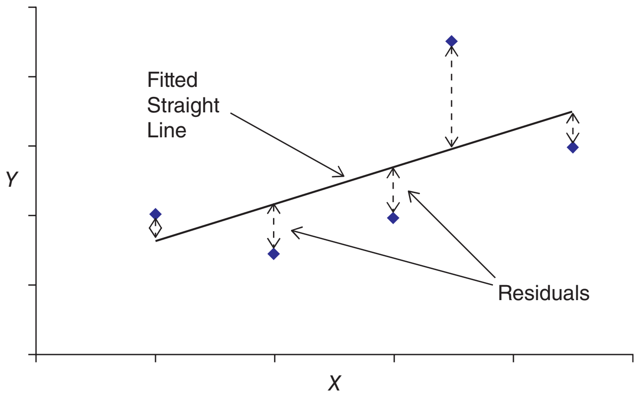


Figure 3.14 Simple linear regression

environment allows managers to supplement these quantitative measures with qualitative estimates.

Linear regression is a parametric method, i.e. the regression model is defined in terms of a finite number of unknown parameters estimated from a set of data. Linear regression involves using the method of Ordinary Least Squares (OLS) to predict the value of a dependent variable Y, based on an independent variable X.[19] More technically, simple linear regression fits a straight line through a set of n data points such that the sum of squared errors, i.e. the vertical distances between the data points and fitted straight line, are as small as possible (see Figure 3.14).

A linear regression model is given by:

$$Y_i = a + bX_i + e_i \tag{3.21}$$

where X_i are the independent[20] variables, Y_i are the dependent[21] variables, b the slope (or gradient) of the straight line, a the intercept and e_i are the error terms.[22]

e_i capture all of the other factors that influence the dependent variables Y_i other than the independent variable X_i. It is a necessary condition that the independent variables X_i are linearly independent, i.e. it is not possible to express any independent variable as a linear combination of the others. This can also be expressed in terms of the regression model having no multi-collinearity, i.e. there is no strong correlation between two or more independent variables. The dependent variables Y_i should also be approximately normally distributed.

[19] For each data point both X and Y are known.
[20] X_i are also known as the regressor, exogenous, explanatory, input or predictor variables.
[21] Y_i are also known as the regressand, endogenous, response or measured variable.
[22] e_i are also known as the residual, disturbance or noise term.

3.9.1 Coefficient of Determination

The strength of the relationship between the dependent and independent variables can be determined by calculating the coefficient of determination (denoted by R^2). R^2 can be interpreted as the proportion of the variance in the dependent variable that is predicted (or explained) by the independent variable. R^2 ranges between zero and one, zero indicating that the dependent variable cannot be predicted from the independent variable and one indicating that there is no error in the relationship between the dependent and independent variable. An R^2 of 0.40 means that 40% of the variance in the dependent variable Y is predicted or explained by the independent variable X. Mathematically, R^2 is the square of the correlation. A related measure to the coefficient of determination is the standard error (SE) in the regression line. Standard error is the average amount by which the regression model over- or under-predicts (or explains). The higher R^2, the lower the standard error, and the more accurate any predictions based on the regression model will be.

3.9.2 Residual Plots

The error term in the linear regression model must possess the following properties:

1. The errors in the regression model must be random variables and have a mean of zero.
2. The variance of the errors must be constant, i.e. *homoscedastic*.[23]

Once a regression model has been fitted to a set of data, a further investigation of the error terms allows another test of the validity of the linearity assumption. The error terms are determined from the difference between the observed values and those predicted from the linear regression model, that is:

$$e = Y - \hat{Y} \tag{3.22}$$

Where $\hat{Y} = a + bX$. A residual plot as shown in Figure 3.15 can reveal the possibility of any nonlinear relationship between the variables as well as the presence of outliers. Such outliers may represent erroneous data, or indicate a poorly fitting regression model. If the residual plot shows a random scatter of data points, as in Figure 3.15, then this is consistent with the model being linear.

The S&P 500 was used to compare the returns for a Long-Short hedge fund index to that of the equity market. The analysis involved regressing monthly returns for the Long-Short Index (dependent variable) against those of the benchmark S&P 500 (independent variable). The regression was run over the period 2008 to 2013. In Source 3.13 the output shows only the statistics that are of interest to the analysis here. R is the correlation coefficient and R^2 is simply

[23] If this condition is violated the errors are said to *heteroscedastic*.

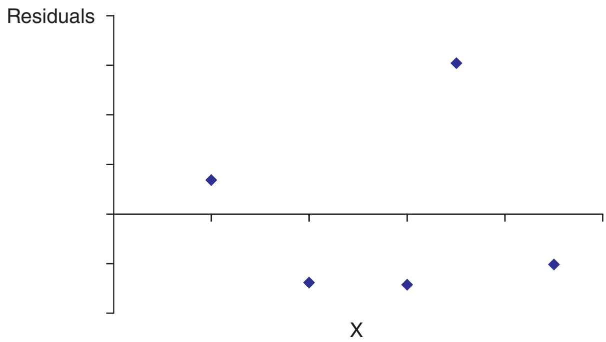


Figure 3.15 A residual plot

```
%File: regression.m
%import XL data
[~,~,returns] = getXLData('hfma_matlab_data.xlsx','LS Index & S&P 500');

x = returns(:,2); %S&P 500 (%)
y = returns(:,1); %Long-Short Index (%)

list = {'beta','adjrsquare','rsquare','tstat'}; %requested statistics
stats = regstats(y,x,'linear',list); %perform regression

%output statistics
cfs = stats.beta; %regression coefficients
arsq = stats.adjrsquare; %adjusted R square
rsq = stats.rsquare; %R square
t = stats.tstat.t; %t-statistic
pv = stats.tstat.pval; %p-value

%other statistics
se = cfs./t; %standard error
correl = corr(x,y); %correlation

OUTPUT:
```

Correlation	0.4012
adjusted R2	0.1490
R2	0.1609

	Coefficients	SE	t-statistic	p-Value
Intercept (a)	0.6086	0.2027	3.0025	0.0037
Slope (b) [S&P500]	0.2178	0.0594	3.6642	0.0005

Source 3.13 MATLAB® code for the linear regression

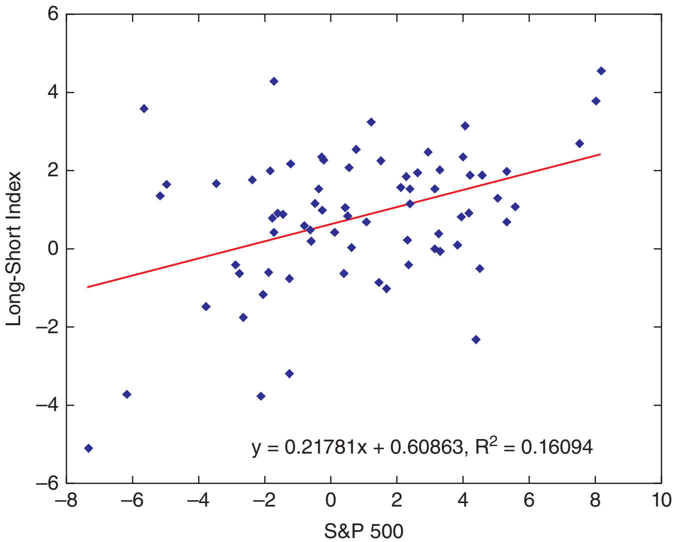


Figure 3.16 The linear regression plot and model

the correlation coefficient squared, i.e. coefficient of determination. The adjR^2 is a version of R^2 that has been adjusted to take into account the number of independent variables in the regression model. R^2 tends to overestimate the strength of the relationship between the dependent and independent variable, especially when there is more than one independent variable. In our case, a relatively high 16.09% of the variance in the Long-Short Index is predicted (or explained) by that in the S&P 500 benchmark. Such a result is likely since the hedge fund strategy is highly equity related, and they will react similarly to the effect of a variety market factors, in particular those that affect the equity markets.

Figure 3.16 shows the plot of the monthly returns for the Long-Short index and S&P 500 as well as the equation of the fitted regression line and R^2 value.

In order to be confident of the validity of such a regression model it is necessary to investigate the statistical significance of the estimated coefficients, especially the slope i.e. $b = 0.2178$ (see Box 3.3).

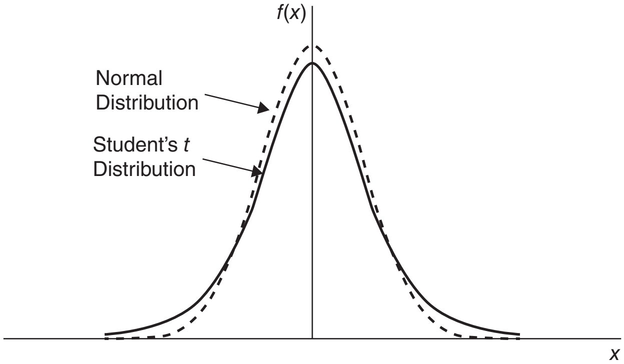
Box 3.3 Statistical Significance

The standard error (SE) about the regression line is a measure of the average amount that the regression equation over- or under-predicts the model. The SE is a measure of the standard deviation of the coefficient in the regression model. The SE is used for calculating the t-statistics used in statistical tests of significance. The t-statistic (t) is given by:

$$t = \frac{b}{SE}$$

Where b is a coefficient calculated from the regression analysis, e.g. the slope. The t-statistic is compared with the value from the Student's t distribution so as to determine a p-Value.

The Student's t-distribution is generally defined as the probability distribution of a set of random variables that best fit the data without knowing the population standard deviation. The particular form of the t-distribution depends on the number of degrees of freedom, i.e. the number of independent observations in a set of data. The higher the degrees of freedom, the closer the t-distribution to the standard normal distribution. The t-distribution is very similar in shape to the normal distribution but has fatter tails resulting in more values being further away from the mean value as shown below.



The t-distribution plays a central role in the associated t-test for assessing the statistical significance of the difference between two sample means, the development of confidence intervals to determine the difference between two population means, and in linear regression analysis.

The p-value is the probability of obtaining a test statistic (e.g. t-statistic) at least as extreme as the one that was actually observed, assuming that the NULL hypothesis is true. It is often necessary to reject the NULL hypothesis when the p-value is less than 0.05, i.e. a 5% chance of rejecting the NULL hypothesis when it is true. When the NULL hypothesis is rejected, the result is said to be statistically significant. That is, if 95% of the t-distribution is closer to the mean than the t-statistic for the coefficient of the regression, this relates to a p-Value of 5%. With such a value there is only a 5% chance that the results of the regression analysis would have occurred in a random distribution, or, there is a 95% probability that the coefficient is having some effect on the regression model. It is important to note that the size of the p-Value for a coefficient indicates nothing about the effect the coefficient is having on the regression model, i.e. it is possible to have a coefficient with a very low p-Value which has only a minimal effect on the model.

```
%File: regression.m
%plot linear regression
p = polyfit(x,y,1);
f = polyval(p,x);

figure()
plot(x,y,'.');
hold on
%linear fit
plot(x,f,'-r');
xlabel('S&P500 (%)');
ylabel('Long-Short Index (%)');
%add linear equation & R-squared
text(-1,-5,['y= ',num2str(cfs(2)),'x + ',num2str(cfs(1)),',
R^2= ',num2str(rsq)]);
```

Source 3.14 MATLAB® code for the linear regression plot in Figure 3.16

If there is a significant relationship between the dependent and independent variable (i.e. Long-Short Index and S&P 500, respectively), the slope (b) will not equal zero. In this case, the NULL hypothesis states that the slope is equal to zero. From Source 3.13 the t-statistic for the slope is 3.6642 which gives a p-Value of 0.0005. The p-Value is the probability that a t-statistic having 70 degrees of freedom is more extreme than 3.6642. Since the p-Value is less than the significance level of 5% (0.05) the NULL hypothesis must be rejected.

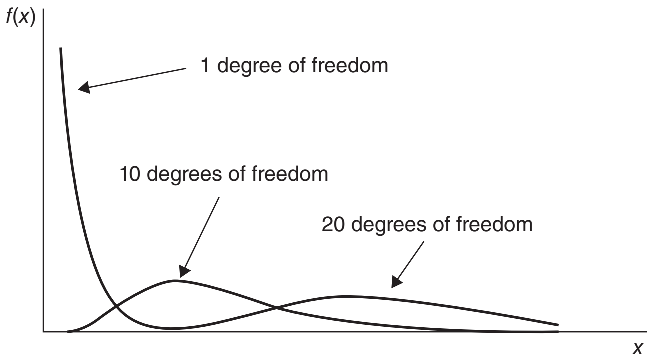
3.9.3 Jarque-Bera Test

We have already looked at several visual methods for testing for normality of a distribution of monthly hedge fund returns, namely inspection and the normal probability plot. Jarque and Bera (1987) developed a statistical test for normality based on the sample skew and kurtosis, namely the Jarque-Bera test. The JB value has a Chi-square distribution (see Box 3.4) with two degrees of freedom (one for the skew and one for the kurtosis) used to test the validity that a distribution of hedge fund returns is normally distributed. More formally, the NULL hypothesis[24] is a joint hypothesis that the skew is zero and kurtosis is three (or excess kurtosis is zero) which are consistent with a normal distribution.

[24] The NULL hypothesis attempts to show that no variation exists between variables, or that a single variable is no different from zero. It is presumed to be true until statistical tests nullify it for an ALTERNATIVE hypothesis. The null hypothesis assumes that any kind of difference or significance you see in a set of data is due only to chance. If a NULL hypothesis is proven true, a p-Value gives the probability that a random sample of data would deviate from the normal distribution as much as the test data.

Box 3.4 Chi-Squared Distribution

The Chi-square (or χ^2) distribution is the sum of the squared values of the number of independent random variables with finite mean and variance drawn from a standard normal distribution. The χ^2 distribution is dependent on the number of degrees of freedom. As the number of degrees of freedom increases the χ^2 distribution approaches the normal distribution as shown below. As a result of the central limit theorem, the χ^2 distribution only converges to a normal distribution only for large values of n, i.e. it is an asymptotic distribution.



The χ^2 distribution is used most commonly in χ^2 tests for the goodness of fit of an observed probability distribution to a theoretical one, and in confidence interval estimation for a population standard deviation.

The JB test statistic is given by:

$$JB = \frac{n}{6}\left(s^2 + \frac{1}{4}k^2\right) \tag{3.23}$$

where n is the sample size, s is the sample skew and k the sample excess kurtosis.

Source 3.15 shows the MATLAB® code for a user-defined function that calculates all the necessary statistics for a Jarque-Bera test of normality. Source 3.16 shows the results of testing for normality for the distribution of monthly hedge fund returns for the hypothetical CTA Index between 2008 and 2013.

For a 95% level of significance, the χ^2 critical value with two degrees of freedom is 5.9915. Using the critical value, the JB test for normality states that:

If *JB value* $>=$ χ^2 *critical value* $\rightarrow$ *reject NULL hypothesis of normality*

```
%File: fJBTest.m
%JARQUE-BERA TEST FOR NORMALITY
function [jb, critval, pval] = fJBTest(x,sig,dof)
%-----------------------------------------------------------------
%x:          returns vector
%sig         significance level e.g. 95% (0.95)
%dof         degrees of freedom e.g. 2
%jb:         Jarque-Bera value
%critval     inverse Chi-square at required significance level
%pval        Chi-square at Jarque-Bera value
%-----------------------------------------------------------------

    n = size(x,1); %# data points
    S = skewness(x,0); %sample skew
    K = kurtosis(x,0)-3; %sample excess kurtosis
    jb = (n/6)*(S^2+(K^2/4)); %Jarque-Bera value
    critval = chi2inv(sig,dof); %inverse Chi-square value
    pval = chi2pdf(jb,dof); %Chi-square at Jarque-Bera value
end
```

Source 3.15 User-defined MATLAB® function for performing a Jarque-Bera test for normality

For a 95% level of significance, using the p-Value (0.3711), the JB test for normality states that:

$$\text{If } p\text{-Value} < 0.05 \rightarrow \textit{reject NULL hypothesis of normality}$$

Source 3.16 shows that the NULL hypothesis of normality cannot be rejected at the 95% level of significance with both the critical and p values, i.e. the distribution of monthly returns for the hypothetical CTA Index between 2008 and 2013 is assumed to be normally distributed.[25] Note that the JB test is only assumed valid for large sample sizes ($n \geq 50$), and can produce misleading results for smaller samples due to the asymptotic nature of the χ^2 distribution and related critical values. The D'Agostino-Pearson test for normality, which is also based on the values of skew and kurtosis, usually gives more accurate results for smaller sample sizes.[26]

In this chapter we have covered the main visual and mathematical techniques for analysing the time series of hedge fund returns. We have noted that such hedge fund returns tend to show departures from the normal distribution and

[25] A Type I error occurs when a NULL hypothesis is rejected when it is in fact true.
[26] Other statistical tests for normality are available, such as the Kolmogorov-Smirnov, Shapiro-Wilks and Anderson-Darling tests. All of these tests have their advantages and disadvantages.

```
%JARQUE-BERA test for normality
%import XL data
[~,~,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

sig = 0.95; %significance level (95%)
dof = 2; %degrees of freedom
[jb,critval,pval] = fJBTest(returns,sig,dof); %get Jarque-Bera test
statistics

OUTPUT:
```

Jarque-Bera Value	0.5960
Sig. Level	95.0000
Degrees of Freedom	2.0000
Critical Value	5.9915
p-Value	0.3711

Source 3.16 MATLAB® code of a Jarque-Bera test for normality

can possess an amount of skewness and kurtosis in their distribution. We also know that hedge fund managers are heavily exposed to the risks involved when using various underlying instruments in their investment strategies. In this case, for a certain hedge fund strategy or index, a set of correlated market factors can be identified and the strength of their relationship quantified using regression analysis. Indeed, along with a thorough knowledge of the particular strategy and market environment, hedge fund managers can further supplement these quantitative measures with qualitative estimates.

4
Mean-Variance Optimisation

Hedge fund analysis primarily involves ascertaining relevant statistical properties of the hedge fund returns distribution in order to make informed decisions about the characteristics and performance of a hedge fund. We have already looked in detail at the two most prominent statistical parameters most often cited: the mean and standard deviation. In 1952, H.M. Markowitz introduced the topic of modern portfolio theory which opened up the possibility of being able to optimise a portfolio of assets so as to minimise the portfolio risk for an acceptable level of portfolio return. Clearly, such a technique is extremely valuable to hedge fund managers, especially when dealing with asset allocation and the efficient distribution of wealth across a portfolio.

Chapter 4 introduces the main ideas behind mean-variance optimisation and shows how a simple hedge fund optimisation problem to minimise the portfolio variance can be set up and solved. The chapter also looks at a common modification of the minimum variance optimisation problem in terms of maximising the Sharpe ratio.

4.1 PORTFOLIO THEORY

4.1.1 Mean-Variance Analysis

All hedge fund managers would like to achieve the highest possible return from their investment portfolios; however, this has to be weighed up against the amount of risk they are willing to accept. Figure 4.1 shows the risk-return scatter plot for the 10 hypothetical hedge funds with monthly returns.

It is not unknown that assets with higher returns generally correlate with higher risk. However, a hedge fund manager can reduce their overall exposure to the risk from individual assets by investing in a well-diversified portfolio of uncorrelated assets. For example, it is generally accepted that equity markets move independently to the bond market and so a combination of both assets in a portfolio can lead to an overall lower level of risk. Indeed, diversification can lead to a reduction in risk even if asset returns are not negatively correlated. By holding a combination of assets that are not perfectly correlated (i.e. $-1 \leq \rho_{ij} < 1$) one can effectively achieve the same expected portfolio return but with a much lower level of portfolio risk (as measured by the portfolio variance). H.M. Markowitz[1] (1952) was one of the first to look at the correlation between

[1] Harry Markowitz (1927–) is an American economist and best known for his pioneering work in Modern Portfolio Theory (MPT), studying the effects of asset risk, return, correlation and diversification on portfolio returns.

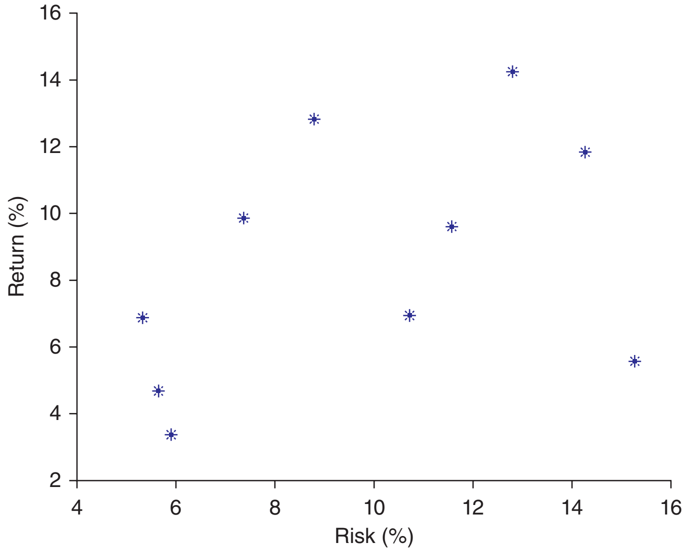


Figure 4.1 Risk-return scatter plot for 10 hypothetical hedge funds

```
%File: riskreturn.m
%import XL data
[names,~,returns] = getXLData('hfma_matlab_data','10 Hedge Funds');

am1 = fMean(returns,12); %annual mean
am2 = fStd(returns,0,12); %annual standard deviation

scatter(am2,am1,'b*') %scatter plot
xlabel('Risk (%)');
ylabel('Return (%)');
```

Source 4.1 MATLAB® code for scatter plot in Figure 4.1

various assets in order to obtain a mean portfolio return and subsequent reduction in the overall risk through diversification.[2] More technically, for a portfolio made up of N risky assets, $i = 1, \ldots, N$, the expected portfolio return r_p is given by:

$$r_p = \sum_{i=1}^{N} w_i \bar{r}_i \tag{4.1}$$

where $\bar{r}_i$ are the mean return associated with each risky asset, i and w_i are the individual holdings (or weights) invested in each risky asset i.

[2] Under the assumption that asset returns are normally distributed, investors are rational and markets are efficient.

It is assumed that the portfolio is fully invested so that the total holdings in each asset always add up to 100%, i.e. the following condition on the weights must hold:

$$\sum_{i=1}^{N} w_i = 1 \tag{4.2}$$

The portfolio variance σ_p^2 is given by:

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_i \sigma_j \rho_{ij} \tag{4.3}$$

where w_i and w_j are the weights, σ_i and σ_j are the standard deviations, and ρ_{ij} is the correlation between the returns of assets i and j.

Note that for any pair of random variables, X_i and X_j, the correlation ρ_{ij} can be written as:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \tag{4.4}$$

where $\sigma_{ij} = \text{cov}(X_i, X_j)$.

The portfolio variance can therefore be written as:

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{ij} \tag{4.5}$$

where, more formally, $\sigma_{ij} = \text{cov}(\bar{r}_i, \bar{r}_j)$ and $\sigma_i^2 = \text{cov}(\bar{r}_i, \bar{r}_i)$.

Both the mean portfolio return and variance can also be transformed into a compact matrix notation. That is, for the mean portfolio return we can write:

$$r_p = \sum_{i=1}^{N} w_i \bar{r}_i = W^T \times R \tag{4.6}$$

where W^T is the matrix transpose of the vector W which contains all the individual asset weights, w_i and R is the vector of mean returns for assets i.

W, W^T and R are given by:

$$W = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ . \\ . \\ w_N \end{pmatrix} \tag{4.7}$$

$$W^T = (w_1 \quad w_2 \quad w_3 \quad . \quad . \quad w_N) \tag{4.8}$$

$$R = \begin{pmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \bar{r}_3 \\ . \\ . \\ \bar{r}_N \end{pmatrix} \tag{4.9}$$

Thus, for the portfolio variance we can write:

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{ij} = W^T \Sigma W \tag{4.10}$$

Where Σ is the variance-covariance matrix[3] for the individual assets i and j given by:[4]

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \mathrm{cov}_{12} & \mathrm{cov}_{13} & . & . & \mathrm{cov}_{1n} \\ \mathrm{cov}_{21} & \sigma_2^2 & \mathrm{cov}_{23} & . & . & . \\ \mathrm{cov}_{31} & \mathrm{cov}_{32} & \sigma_3^2 & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \mathrm{cov}_{n1} & . & . & . & . & \sigma_N^2 \end{pmatrix} \tag{4.11}$$

Markowitz's work ultimately led to the development of mean-variance portfolio optimisation[5] as a method of achieving a desired level of portfolio expected return for a degree of portfolio risk. Given a target expected return of r*, the mean-variance optimisation problem can be stated as follows:

$$\min W^T \Sigma W$$

Subject to the constraints:

$$W^T \times R = r^*$$

$$\sum_{i=1}^{N} w_i = 1$$

[3] This is also known as the VCV matrix.

[4] As with all covariance matrices, Σ must be positive definite.

[5] Such a technique is known as a Quadratic Programming (QP) problem involving the optimisation (either minimising or maximising) of a quadratic function of several variables subject to a set of linear constraints on these variables. In this case, portfolio variance is a quadratic function of the weights w_i.

The above optimisation problem assumes that short selling is allowed, i.e. the weights can be negative. If no short selling is allowed, the optimisation problem can be modified by adding another constraint:

$$\min W^T \Sigma W$$

Subject to the constraints:

$$W^T \times R = r^*$$

$$\sum_{i=1}^{N} w_i = 1$$

$$w_i \geq 0$$

It is also possible to impose a constraint on the maximum allowable investment limits (or bounds) on one or more of the assets in the portfolio, i.e. $w_i \leq b$ where b is a real number.

4.1.2 An Optimisation Problem

A mean-variance portfolio optimisation problem can be solved in MATLAB® using the built-in MATLAB® function `fmincon()` which finds the minimum of a problem specified by:

$$\min_x f(x) \text{ such that } \begin{cases} c(x) \leq 0 \\ ceq(x) = 0 \\ A \cdot x \leq b \\ Aeq \cdot x = beq \\ lb \leq x \leq ub \end{cases}$$

where b and beq are vectors, A and Aeq are matrices making up the linear equality and inequality constraints, and $f(x)$ is a function that returns a scalar. $c(x)$ and $ceq(x)$ are vectors that can be nonlinear functions and generally used for more challenging optimisation problems. lb and ub represent vectors for the upper and lower bound limits, respectively.

`fmincon` attempts to find a constrained minimum of a scalar function of several variables starting at an initial estimate. This is generally referred to as constrained nonlinear optimisation or nonlinear programming. The general form of the setup of `fmincon` is given as:

```
[x,fval,exitflag] = fmincon(fun,x0,A,b,Aeq,beq,lb,ub,[],options)
```

The optimisation starts at `x0` and attempts to find a minimiser `x` of the objective function described in `fun` subject to the linear inequalities `Aeq *x = beq`

```
%File: optimisation.m
%import XL data
[names,~,returns] = getXLData('hfma_matlab_data.xlsx','10 Hedge Funds');

N = size(returns,2); %size of portfolio (N)
R = fMean(returns,12)'; %annual returns (R)
W0 = repmat(1/N,N,1); %initial weights (W0)

covar = cov(returns); %covariance matrix

for i=1:N
  for j=1:N
      if i == j
          covar(i,j)=covar(i,j)*12; %annualise variance for monthly returns

      end
  end
end
%VCV matrix
VCV = covar;

pR = W0'*R; %annual portfolio return
pVar  = W0'*VCV*W0; %annual portfolio variance

OUTPUT:
```

	Annual Returns (%)
CTA1	12.823
CTA2	6.860
CTA3	14.211
GM1	11.810
GM2	5.522
LS1	9.842
LS2	9.573
LS3	6.913
MN1	4.646
MN2	3.322

	Initial Weights (W0)
CTA1	0.10
CTA2	0.10
CTA3	0.10
GM1	0.10
GM2	0.10
LS1	0.10
LS2	0.10
LS3	0.10
MN1	0.10
MN2	0.10

	Annual (%)
Port. Mean	8.552
Port. Variance	10.726

	CTA1	CTA2	CTA3	GM1	GM2	LS1	LS2	LS3	MN1	MN2
CTA1	77.414	-0.279	0.160	-0.079	-1.835	-0.242	0.377	1.079	0.451	1.152
CTA2	-0.279	28.215	-1.390	0.337	0.175	-0.225	0.521	0.031	-0.184	-0.791
CTA3	0.160	-1.390	163.258	0.649	-0.081	-0.461	-1.218	0.659	1.549	0.555
GM1	-0.079	0.337	0.649	202.696	0.192	0.389	-1.487	-1.048	1.346	-0.453
GM2	-1.835	0.175	-0.081	0.192	233.159	-0.671	0.300	-1.790	-0.095	-1.746
LS1	-0.242	-0.225	-0.461	0.389	-0.671	54.225	1.370	-0.602	0.115	-0.045
LS2	0.377	0.521	-1.218	-1.487	0.300	1.370	134.236	0.961	-0.487	0.476
LS3	1.079	0.031	0.659	-1.048	-1.790	-0.602	0.961	114.905	0.682	0.774
MN1	0.451	-0.184	1.549	1.346	-0.095	0.115	-0.487	0.682	31.815	-0.140
MN2	1.152	-0.791	0.555	-0.453	-1.746	-0.045	0.476	0.774	-0.140	34.795

Source 4.2 MATLAB® code for the calculation of the annual portfolio return and variance for an equally weighted portfolio

and `A *x` $\leq$`b` as well as defining a set of lower and upper bounds on the design variables in `x`, so that the solution is always in the range `lb`$\leq$`x`$\leq$`ub`. `exitflag` is an identifier for the reason the algorithm terminated. `fmincon` uses one of four algorithms, namely `active-set`, `interior-point`, `SQP`,[6] and `trust-region-reflective` all of which have their particular use depending on the optimisation problem. The required optimisation algorithm can be set in `options` as can a range of other parameters, for example maximum number of iterations and maximum allowable tolerance, etc. `[]` is used when that particular parameter is not required for the optimisation.[7]

Before using `fmincon` we must first set up the initial problem which involves determining the annual portfolio return and variance for an equally weighted portfolio. The portfolio will comprise the 10 hypothetical hedge funds of monthly returns previously introduced in Chapter 3. Source 4.2 shows the calculation of the Σ (i.e. VCV) and R matrices required to determine the portfolio return and variance using an initial equal weighting of 0.10 for each of the 10 hypothetical hedge funds. The annual portfolio return and variance for the 10 hypothetical hedge fund portfolio are 8.552% and 10.726% (i.e. σ = 3.275%), respectively.

The next step is to set up the necessary linear constraints and upper and lower bounds for the optimisation problem. In this case we will be restricting our analysis to one of no short sales, lower bounds greater than or equal to zero with an upper bound of 0.50, and a target expected return[8] of 9.00% (see Source 4.3). Box 4.1 gives an overview of how to handle various types of constraints when transforming them into vectors and matrices.

Box 4.1 Constraints

Bound Constraints

Bound constraints represent the lower and upper bounds on individual components, such that:

$$x \geq l \text{ and } x \leq u$$

Lower and upper bounds limit the components of the solution x. Generally, if the bounds on the location of an optimum are known, the solution can

[6] Sequential Quadratic Programming (SQP).

[7] For further details please refer to the *Optimisation Toolbox User's Guide* available from MathWorks.

[8] Clearly, the targeted expected return cannot be more than the highest return for each asset; unless negative holdings are possible, i.e. short sales are allowed.

be obtained much faster and be more reliable if these bounds are explicitly including problem formulation. Bounds should be given as vectors with the same length as *x*, or as matrices with the same number of elements as *x*.

Linear Inequality Constraints

Linear inequality constraints are written in the form:

$$\mathrm{A} \cdot \mathrm{x} \leq= \mathrm{b}$$

Where *A* is an *m*-by-*n* matrix, which represents *m* constraints for an *n*-dimensional vector *x*. *b* is also *m*-dimensional. Note that "greater than" constraints can be written as "less than" constraints by multiplying through by –1, for example a constraint of the form $A \cdot x \geq b$ is equivalent to the constraint $-A \cdot x \leq -b$.

Suppose we have the following linear inequalities as constraints:

$$\begin{aligned} &x_1 + x_3 \leq 4 \\ &2x_2 - x_3 \geq -2 \\ &x_1 - x_2 + x_3 - x_4 \geq 9 \end{aligned}$$

Here $m = 3$ and $n = 4$. These can be written using the following matrix *A* and vector *b* representations:

$$\mathrm{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -2 & 1 & 0 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

$$b = \begin{bmatrix} 4 \\ 2 \\ -9 \end{bmatrix}$$

Note that the "greater than" inequalities have first been multiplied by –1 to transform them into "less than" inequalities. In MATLAB® these would be written as:

```
A = [1 0 1 0;
      0 -2 1 0;
     -1 1 -1 1];
b = [4;2;-9];
```

Linear Equality Constraints

Linear equalities are written in the form:

$$Aeq \cdot x = beq$$

Where *Aeq* is an *m*-by-*n* matrix, which represents *m* constraints for an *n*-dimensional vector *beq*.

Note that it is sometimes possible to write constraints in several ways, for example, with a constraint $5\ x \leq 20$, write it as a bound $x \leq 4$ instead of a linear inequality or nonlinear inequality. That is, for best optimisation results always use the lowest numbered constraints possible.

Source 4.4 shows the objective function that is used in the minimum variance optimisation problem shown in Source 4.5. The optimised weights (*W*) give a minimum annual portfolio variance of 8.407% (i.e. $\sigma = 2.899\%$) which is less than the original value of 10.726% (i.e. $\sigma = 3.275\%$). That is, the portfolio of weighted assets has been optimised to give the minimum variance (or risk) for a higher expected portfolio return. Clearly, as market conditions and risk preferences change, the portfolio must be rebalanced and a new set of optimised weights determined. In the original work by Markowitz, the cost of the transactions associated with buying and selling assets was not included. However, such transaction costs could represent a significant amount to the hedge fund manager and need to be taken into account when considering mean-variance optimisation problems.

4.1.3 Sharpe Ratio Maximisation

Rather than minimising the variance of portfolio returns it is often the case that practitioners will instead maximise the Sharpe ratio since it is a much more insightful performance metric. Although the Sharpe ratio will be covered in a later chapter we include it here so as to develop an optimisation based on the Sharpe ratio. The Sharpe ratio[9] was first introduced by Sharpe (1994) in the following form:

$$Sharpe = \frac{R_P - r_f}{\sigma_P} \tag{4.12}$$

where R_P is the annual mean return, r_f is the annual risk-free rate and σ_P is the annual standard deviation (or volatility) of the returns.

[9] William Sharpe (1934–) is famous for the development of the Capital Asset Pricing Model (CAPM pronounced *Cap-Em*).

```
%File: optimisation.m
%<....>

%target return, linear constraints and bounds
rstar = 9.0; %target return
Aeq = [R';ones(1,N)];
beq = [rstar;1];
lb = zeros(N,1);
ub = ones(N,1)*0.5;

OUTPUT:

Aeq =

  12.8234 6.8600 14.2111 11.8096 5.5221 9.8420 9.5734 6.9130 4.6458 3.3217
   1.0000 1.0000  1.0000  1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000

beq =

    9
    1

lb =

    0
    0
    0
    0
    0
    0
    0
    0
    0
    0

ub =

   0.5000
   0.5000
   0.5000
   0.5000
   0.5000
   0.5000
   0.5000
   0.5000
   0.5000
   0.5000
```

Source 4.3 MATLAB® code for the implementation of the linear constraints, and upper and lower bounds

```
%File: minVar.m
%OBJECTIVE FUNCTION: MINIMINIMISE VARIANCE
function fval = minVar(W,VCV)

%-------------------------------------------------------------------

%INPUTS:
%W      - weight vector
%VCV    - VCV matrix

%-------------------------------------------------------------------

  fval = W'*VCV*W;
end
```

Source 4.4 MATLAB® code for the objective function for minimising portfolio variance

The Sharpe ratio introduces the concept of a static benchmark to the numerator. In other words, the investor wants to earn at least the risk-free rate and since they want the highest Sharpe ratio, this will begin to penalise hedge funds whose returns are low. Also, any hedge fund with annualised returns less than the risk-free rate will have a negative Sharpe ratio regardless of the hedge fund volatility.

Source 4.6 shows the results for the initialisation of the Sharpe ratio maximisation problem using 10 hypothetical hedge funds with daily returns for an equally weighted portfolio. The annual risk-free rate, r_f, has been set to 2.0%.

As with the minimisation of variance, the next step is to set up the necessary linear constraints and upper and lower bounds for the optimisation problem. Again, we will be restricting our analysis to one of no short sales, lower bounds greater than or equal to zero and with an upper bound of 0.50. (see Source 4.7).

Source 4.8 shows the objective function that is used in the maximisation of the Sharpe ratio optimisation problem shown in Source 4.9. The optimised weights (W) give an annual Sharpe ratio of 2.851% which is greater than the original value of 2.687%.

4.2 EFFICIENT PORTFOLIOS

Markowitz reduced the mean-variance optimisation problem to that of finding the locus efficient portfolios, with "efficient" points having the highest expected portfolio return for a given level of portfolio variance (i.e. risk). Portfolios

```
%File: optimisation.m
%<....>

%minimum variance optimisation
options = optimset('LargeScale', 'off', 'Algorithm','sqp');
[W,fval,ExitFlag] = fmincon(@(W)
minVar(W,VCV),W0,[],[],Aeq,beq,lb,ub,[],options); %calls minVar.m
  disp('Minimum Variance');
  if (ExitFlag ~= 1)
      disp('Optimization did not converge!');
  else
      disp('Optimization converged');
  end

OUTPUT:
```

	Optimised Weights (W)
CTA1	0.163
CTA2	0.221
CTA3	0.088
GM1	0.056
GM2	0.022
LS1	0.173
LS2	0.066
LS3	0.051
MN1	0.103
MN2	0.057

	Minimised (%)
Port. Variance	8.407

Source 4.5 MATLAB® code for the calculation of the optimised weights (*W*) and minimum portfolio variance for the mean-variance optimisation problem

```
%File: sharperatio.m
%import XL data
[names,~,returns] = getXLData('hfma_matlab_data.xlsx','10 Hedge Funds II');

N = size(returns,2); %size of portfolio (N)
R = mean(returns)'*252; %annualised mean (R)
W0 = repmat(1/N,N,1); %initial weights (W0)
rf = 2.0; %annual risk-free rate

covar = cov(returns); %covariance matrix

for i=1:N
  for j=1:N
      if i == j
          covar(i,j)=covar(i,j)*252; %annualise variance for daily returns
      end
  end
end

VCV = covar; %VCV matrix

pR = W0'*R; %annual portfolio return
pVar  = W0'*VCV*W0; %annual portfolio variance
pSharpe = (pR - rf)./sqrt(pVar); %annual portfolio Sharpe ratio

OUTPUT:
```

	Monthly Returns (%)
HF1	21.669
HF2	14.902
HF3	9.480
HF4	10.336
HF5	13.443
HF6	18.036
HF7	13.859
HF8	13.898
HF9	16.393
HF10	12.857

	Initial Weights (W0)
HF1	0.10
HF2	0.10
HF3	0.10
HF4	0.10
HF5	0.10
HF6	0.10
HF7	0.10
HF8	0.10
HF9	0.10
HF10	0.10

	Annual (%)
Port. Mean	14.487
Port. Variance	21.597
Port. Sharpe	2.687

	HF1	HF2	HF3	HF4	HF5	HF6	HF7	HF8	HF9	HF10
HF1	305.757	0.015	0.042	-0.067	-0.019	-0.019	0.027	-0.010	0.033	0.015
HF2	0.015	200.371	0.008	0.013	0.007	-0.041	0.003	0.007	0.013	-0.015
HF3	0.042	0.008	126.350	-0.010	-0.002	0.047	-0.006	-0.007	-0.028	-0.008
HF4	-0.067	0.013	-0.010	376.386	0.007	-0.026	0.039	-0.027	-0.004	0.005
HF5	-0.019	0.007	-0.002	0.007	166.627	0.011	-0.001	-0.001	-0.011	0.003
HF6	-0.019	-0.041	0.047	-0.026	0.011	304.742	-0.010	0.041	-0.018	-0.020
HF7	0.027	0.003	-0.006	0.039	-0.001	-0.010	102.758	-0.014	-0.020	-0.004
HF8	-0.010	0.007	-0.007	-0.027	-0.001	0.041	-0.014	164.174	-0.036	0.018
HF9	0.033	0.013	-0.028	-0.004	-0.011	-0.018	-0.020	-0.036	234.915	0.007
HF10	0.015	-0.015	-0.008	0.005	0.003	-0.020	-0.004	0.018	0.007	177.713

Source 4.6 MATLAB® code for the calculation of the annual portfolio return, variance and Sharpe ratio for an equally weighted portfolio

```
%File: sharperatio.m
%<....>

%linear constraints and bounds
A = [ones(1,N);-ones(1,N)];
b = [1;-1];
lb = zeros(N,1);
ub = ones(N,1)*0.5;

OUTPUT:

A =


     1     1     1     1     1     1     1     1     1     1
    -1    -1    -1    -1    -1    -1    -1    -1    -1    -1

b =


     1
    -1

lb =
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0

ub =

    0.5000
    0.5000
    0.5000
    0.5000
    0.5000
    0.5000
    0.5000
    0.5000
    0.5000
    0.5000
```

Source 4.7 MATLAB® code for the implementation of the linear constraints, and upper and lower bounds

```
%File: maxSharpe.m
%OBJECTIVE FUNCTION: MAXIMISE SHARPE RATIO
function fval = maxSharpe(W,R,rf,VCV)

%-------------------------------------------------------------------

%INPUTS:
%W      - weight vector
%R      - returns vector
%rf     - risk-free rate
%VCV    - VCV matrix

%-------------------------------------------------------------------


  R = R-rf;
  fval = -W'*R/sqrt(W'*VCV*W);
end
```

Source 4.8 MATLAB® code for the objective function for maximising the portfolio Sharpe ratio

with the best combinations of weights lie on the efficient frontier, where a target expected return is achieved with minimum risk. The mean-variance efficient frontier contains a unique minimum variance, and overall there will be a minimum variance portfolio. Figure 4.2 shows the efficient frontier for a set of risky assets.

As shown in Figure 4.2, all possible combinations of the risky assets can be plotted in so-called risk-expected return space, and the collection of all such possible portfolios defines a region in this space. The left boundary of this region is a hyperbola with the upper edge being the efficient frontier (also known as the *Markowitz bullet*).

Figure 4.3 shows the efficient frontier for the 10 hypothetical hedge funds with monthly returns. The built-in MATLAB® function `portopt()` returns portfolios on the mean-variance efficient frontier given a specified set of portfolio constraints. If a set of constraints are not specified a set of defaults are used, such as constraints scale the total value of the portfolio to 1, and place the minimum weight of every asset at 0 to prevent short selling. Source 4.10 shows the MATLAB® code for the determination of the efficient frontier.

It is also possible to randomly generate a set of portfolio weights, and combine the random portfolios within an existing mean-variance efficient frontier for comparison as implemented in Source 4.11 and shown in Figure 4.4. In this example we randomly generate the asset weights of 1000 portfolios uniformly distributed on the set of portfolios.

```
%File: sharperatio.m
%<....>

%maximise Sharpe ratio optimisation
options = optimset('LargeScale', 'off', 'Algorithm','sqp');
[W,fval,ExitFlag] = fmincon(@(W)
maxSharpe(W,R,rf,VCV),W0,A,b,[],[],lb,ub,[],options);
disp('Maximise Sharpe');
    if (ExitFlag ~= 1)
        disp('Optimization did not converge!');
    else
        disp('Optimization converged');
    end

OUTPUT:
```

	Optimised Weights (W)
HF1	0.100
HF2	0.100
HF3	0.092
HF4	0.035
HF5	0.107
HF6	0.082
HF7	0.180
HF8	0.113
HF9	0.096
HF10	0.095

	Maximised (%)
Port. Sharpe	2.851

Source 4.9 MATLAB® code for the calculation of the optimised weights (W) and maximised portfolio Sharpe ratio

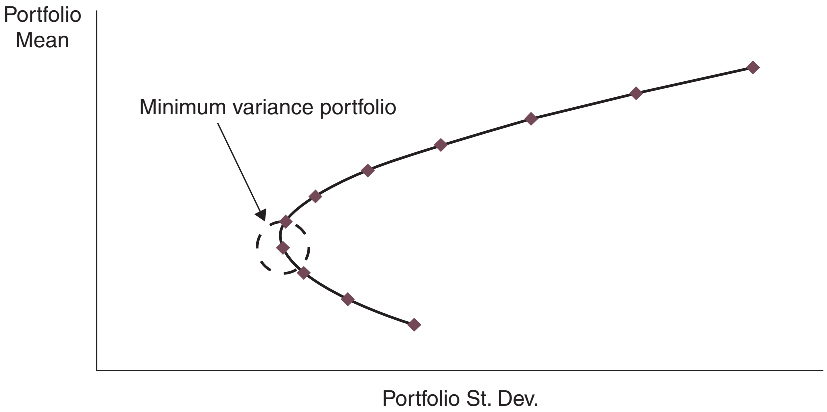


Figure 4.2 The efficient frontier and minimum variance portfolio

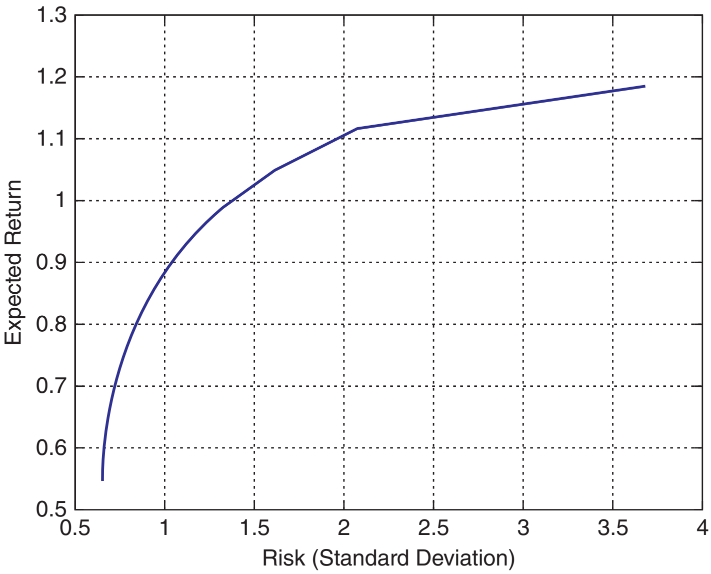


Figure 4.3 Mean-variance efficient frontier

```
%File: efficfron.m
%import XL data
[names,~,returns] = getXLData('hfma_matlab_data','10 Hedge Funds');

N = size(returns,2); %# data points
am1 = fMean(returns,1); %monthly mean
covar = cov(returns); %covariance matrix

portopt(am1,covar,N) %plot mean-variance efficient frontier
```

Source 4.10 MATLAB® code for mean-variance efficient frontier plot in Figure 4.3

```
%File: randport.m
%import XL data
[names,~,returns] = getXLData('hfma_matlab_data','10 Hedge Funds');

N = size(returns,2); %# data points
am1 = fMean(returns,1); %monthly means
covar = cov(returns); %covariance matrix

portopt(am1,covar,N) %plot mean-variance efficient frontier

w = exprnd(1,1000,N); %randomise asset weights (1,000)
tot = sum(w,2);
tot = tot(:,ones(N,1));
w = w./tot;

[pRisk , pRet] = portstats(am1,covar,w); %calculate expected return
&  standard deviation for each portfolio


hold on
plot(pRisk,pRet,'.r'); %plot random portfolios
hold off
```

Source 4.11 MATLAB® code for the random portfolios plot in Figure 4.4

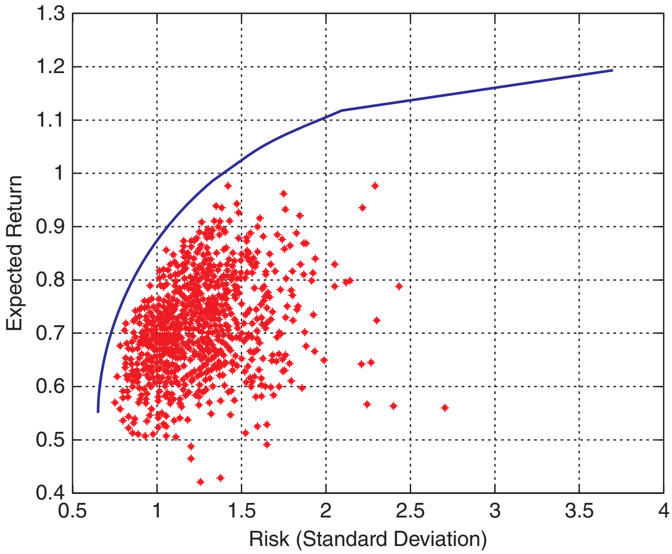


Figure 4.4 Random portfolios within the mean-variance efficient frontier

In this chapter we have introduced the main ideas behind mean-variance optimisation and how an optimisation problem can be set up to either minimise the portfolio variance or maximise the Sharpe ratio. Such techniques can prove invaluable to hedge fund managers who are often tasked with managing the risk of the fund and efficient allocation of resources across the portfolio. Indeed, many hedge fund managers develop increasingly complex optimisation problems utilising a range of performance metrics and risk measures.

5
Performance Measurement

Hedge fund managers and CTAs can use leverage and take short positions (as opposed to traditional managers who cannot), and since their returns over time will be a direct function of leverage and their long-short portfolio mix, so the manager's performance should be measured on a cash basis (unleveraged) "relative" to the portfolio "risk" in order to offset the effects of leverage. The risk proxy measure used is generally based on the second moment of the distribution of returns, e.g. the "volatility" or some other measurable statistical estimate of the variation of the spread of returns associated with the manager. This chapter will look at the various performance measurements which can be used to analyse hedge fund returns in a risk-adjusted sense using the most common metrics applied in industry and academia.

5.1 THE INTUITION BEHIND RISK-ADJUSTED RETURNS

Consider two CTAs – Manager A and Manager B and assume both managers are operating in approximately the same style (e.g. diversified managed futures) and use similar underlying instruments for investment and trading purposes. Also assume that *notional funding*[1] is available for an investor who has \$5 million to collateralise a managed account with their FCM of choice and is hoping one CTA will manage all or part of their capital through a Power Of Attorney (POA)[2] at 2% management fee and 20% incentive fee.[3] Neglecting higher moments, the annualised first two moments (M1 = return and M2 = volatility[4]) and the risk-adjusted return ratio = M1/M2 are shown in Table 5.1. It is assumed in this example that the return data for manager A has been

[1] *Notional funding* can be used to maximise investment capital efficiency since with a margined account, only a certain amount of collateral is required for the margin (M). Usually the investor adds a cushion (C) of capital for future potential drawdowns. If the managed account has a minimum acceptable size of P, then actual funds committed are $A = M + C$, and the notional funding amount is $N = P - A$. It follows that the leverage G obtained by the investment as a function of notional funding for a fixed minimum account size is given by $G = P/A$ to one.

[2] The PoA is an official document signed by the CTA, the investor and the FCM which legally authorises the CTA to manage the client's account at the FCM and receive payment for their management services.

[3] 2% of AuM and 20% of net new profits is the industry standard fee structure for CTAs.

[4] Strictly speaking the second moment of a normal distribution is the *variance*.

Table 5.1 Two CTAs: Managers A and B with Excel and MATLAB® equivalent functions

	A	B	C
1	Month	Mgr A	Mgr B
2	Jan-05	0.10	0.89
3	Feb-05	1.79	-0.90
4	Mar-05	2.04	-5.38
5	Apr-05	1.34	1.33
6	May-05	-0.16	0.88
7	Jun-05	-1.40	-1.02
8	Jul-05	1.15	-2.88
9	Aug-05	-1.20	3.81
10	Sep-05	-1.99	5.93
70	Sep-10	1.06	5.06
71	Oct-10	-0.69	6.19
72	Nov-10	1.27	-4.46
73	Dec-10	0.10	0.59

	D	E	F	G
2		Mgr A	Mgr B	
3	M1	6.86	11.81	=AVERAGE(C2:C73)*12
4	M2	5.27	14.14	=STDEVP(C2:C73)*SQRT(12)
5	M1/M2	1.30	0.84	=F3/F4

Command Window

```
>> muA=mean(mgrA)*12

muA =

    6.8600

>> sigmaA=std(mgrA,1)*sqrt(12)

sigmaA =

    5.2747

>> M1M2ratioA=muA/sigmaA

M1M2ratioA =

    1.3005
```

Command Window

```
>> muB=mean(mgrB)*12

muB =

   11.8096

>> sigmaB=std(mgrB,1)*sqrt(12)

sigmaB =

   14.1379

>> M1M2ratioB=muB/sigmaB

M1M2ratioB =

    0.8353
```

either imported into MATLAB® or as in this case, created and named within MATLAB® as vectors using Excel-link (mgrA and mgrB being the vector names).

Figure 5.1 shows the VAMI for Manager A and B. At first glance an unsophisticated investor who had to choose between Manager A and B (i.e. an *all-or-nothing* investment choice) would choose Manager B, since Manager B obviously has a historical record of providing higher returns over time for their investors. However due to the possibility of notional funding commonly found in the world of managed account investing[5] with CTAs, a sophisticated investor would notice that he or she could fund each account at a range of funding levels so that it would be up to the CTA to decide at what actual level of funds he would calibrate their limits for trading and investing. As a result, a range of returns could be sought for each CTA regardless of their M1 or annualised return performance record, i.e. notional funding can be used to set an investment return target, regardless of the CTA's historical annualised returns and future expectations (see Example 5.1).

[5] Note that notional funding is usually associated with investing in managed accounts, i.e. hedge funds usually have pre-prescribed risk and return parameters in which it is difficult to leverage up or down.

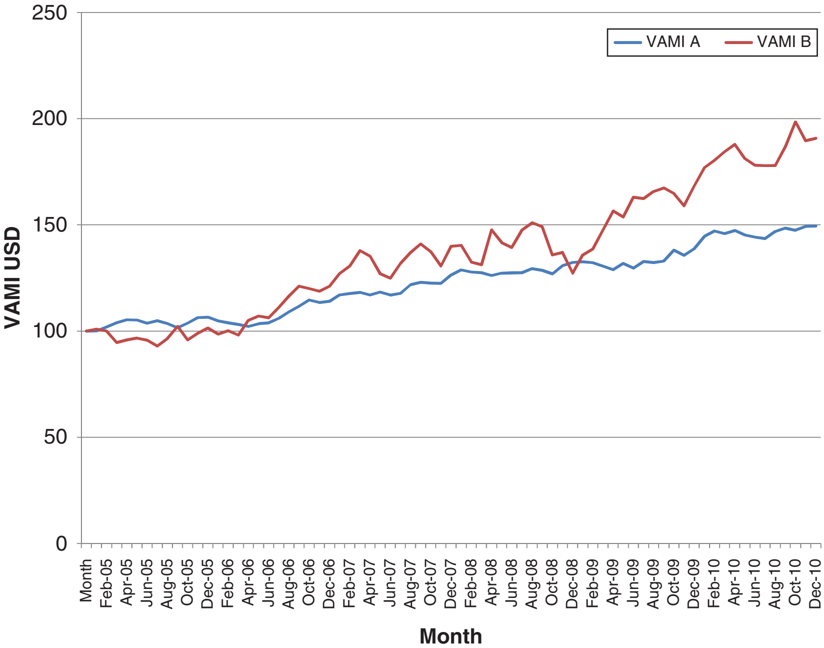


Figure 5.1 VAMI for Two CTAs: Manager A and B

5.1.1 Risk-Adjusted Returns

Due to the notional funding argument discussed in the previous section, it can be seen that it is not just the returns that are important for maximising capital efficiency (as well as performance) of an investment. Assuming for now that only the first two moments are important for investment appraisal, the logical way to proceed for investment analysis is to create a Cartesian[6] diagram in these two dimensions. The concept of investment analysis in *risk-adjusted return*, *mu-sigma* or *mean-variance* space was pioneered by Harry Markowitz (1952).[7] The two dimensional Cartesian points for each CTA are shown in Figure 5.2.

This visualisation method for looking at unique points per manager in two dimensional space is one way to allow the investor to visualise their investment. The major drawback is of course that it assumes only the first two moments of the return distribution are important. However, whilst being close to the truth in reality, it is not quite correct since hedge funds also suffer from higher moment

[6] Descartes (1596–1650) was a French mathematician and philosopher who developed the Cartesian coordinate system.

[7] Nowadays, Markowitz's work is considered to be the foundation of Modern Portfolio Theory (MPT).

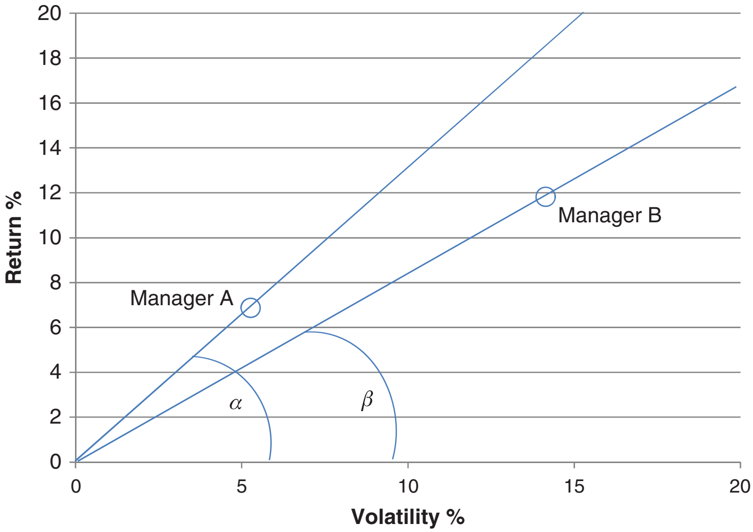


Figure 5.2 CTAs plotted in mu-sigma space

risks as shown by Favre and Ranaldo (2005). For now though, we assume that the investor is only interested in the returns and the riskiness of those returns as measured by volatility. Assuming the investor faces an all-or-nothing choice, which manager would be preferable? Since we know due to notional funding that an investor can in theory leverage up their investment, the objective function for the investor is simply that of a rational investor. For a certain fixed return expectation, they want to arrive at their target wealth expectation with the *minimum of risk* borne in the process. Since we are measuring risk at this stage solely by volatility, it therefore follows that if we were to draw two straight lines from the origin of Figure 5.2 through the unique location point of each manager, and measure the angle of each line from the x-axis, e.g. α and β angles, the manager of preference in the risk-adjusted sense would be the manager who has the largest angle. In this case, it is clear that $\alpha > \beta$, so CTA A is preferred over CTA B in the risk-adjusted sense (since it "dominates" in mu-sigma space it is said to be the more "efficient" of the two portfolios). As can be seen from Figure 5.2, each angle is directly proportional to the *gradient*[8] and it follows that:

$$\frac{R_A}{\sigma_A} > \frac{R_B}{\sigma_B} \tag{5.1}$$

[8] A *gradient* is measured as the ratio of vertical distance to the origin over the horizontal distance to the origin for a point in two-dimensional space.

The investor's risk-adjusted investment objective is usually therefore to maximise future risk-adjusted returns:

$$\theta = \left[\frac{E(R_p)}{\sigma_p}\right]_{\max} \tag{5.2}$$

where $E(R_p)$ is the expected future returns of the investment and σ_p the expected volatility of returns.

It can be seen that the gradient is indeed the same measure as the Sharpe ratio if we neglect the risk-free rate leading to the common industry approach of optimising a portfolio or hedge fund based on its Sharpe ratio as an objective function.

As can be seen from Table 5.1, for the ratio M1/M2, CTA A is the obvious winner with a value of 1.30 and so CTA A is preferred over CTA B in the risk-adjusted sense. Example 5.1 demonstrates how important these concepts are in reality for the investor authorised to use notional funding.

Example 5.1 Notional Funding in Practice

Consider the data for CTA managers A and B above which we will use for this example and assume the results shown in Table 5.1 are net after fees and that the risk-free rate is 0.0% for simplification purposes. You are a young savvy investor with an appetite for high return and have carried out all your research and due diligence and particularly like two CTAs – CTA A and CTA B – with whom you have decided to allocate your money. Since you are also rather busy you don't want to go to the hassle of filling out all the paper work and forms necessary to set up two managed accounts, and instead want to invest all your capital with one CTA with whom you wish to develop a long-term relationship. In particular you want to earn around 30.0% per annum on your capital of \$5m (each manager accepts minimum account sizes of \$1m). Since CTA A has the highest risk-adjusted return ratio you have decided to choose them as your only manager. From their track record, you know that on average you have an expectation that they can produce annualised net returns of 6.86% with an estimated volatility of 5.27% on the minimum account sizes of \$1m. Considering you want to earn around 30.0% per annum and not just 6.86%, and the manager accepts notional funding, this means you need to leverage (or gear up) the managed account by a factor $G = 30.0/6.86 = 4.37$ times. That is, the investor and the CTA should agree to set up an account which is leveraged 4.37 times using notional funding. From Footnote 2, $G = P/A$ and since $G = 4.37$ and $A = \$5\text{m}$, $P = \$21.85\text{m}$. Since $N = P - A$, $N = \$16\text{m}$. Therefore, the investor should agree with the CTA to design an account which should be traded as if it were AuM representing \$21.85m, of which \$5m are the actual funds and the rest are notional funds of \$16.85m. With this gearing, the investor will expect a net return per annum of 30.0% on their \$5m through the CTA. They will, however, also expect an associated higher volatility of G times the CTA volatility, i.e. $4.37 \times 5.27\% = 23.0\%$ since the risk-adjusted ratio of the CTA remains constant under gearing and if the returns (numerator) are increased by a gearing factor G, the volatility (denominator) must also be increased by a factor G for the risk-adjusted return ratio to remain constant.

Since the investor is young and \$5m represents only about 10% of their net worth, they are happy to accept such a level of volatility for an expected annual return of 30.0% on the invested amount of \$5m. Analysis of the manager's disclosure document shows that it is statistically acceptable to use notional funding since the manager claims to have an average *margin-to-equity* (M/E)[9] of 5.0%. With an account size of \$21.85m, this will translate into an operational daily requirement of margin equal to 5.0% of \$21.85m = \$1.09m (= M). Since the investor is investing \$5m in actual funds (A), the cushion $C = A - M = 5.0 - 1.09 = \$3.91$m. Statistically this cushion should be enough to prevent any future margin calls with a low probability of having to inject further collateral into the account in the event of a margin call which could happen due to the manager's poor future performance or an increase in margin levels as set by the futures exchanges or both. The notional funding level is defined as

$$NF = (N \times 100)/P = (16.85 \times 100)/21.85 \approx 77\%.$$

Note the situation if the investor had been uninformed as to the virtues of risk-adjusted investing through CTAs. If they had erroneously chosen CTA B (at first glance perhaps because CTA B had higher returns) to create their 30% return expectation account using notional funding, their volatility expectation would have been higher at 35.9% (30.0/11.81 × 14.14). This example demonstrates that by using risk-adjusted returns to measure the manager's relative performance and by correctly choosing CTA A to invest in, their expected volatility for reaching an annual return of 30.0% would be around 23.0% for CTA A versus 35.9% for CTA B – a significant difference of 12.9%.

This example shows why in the world of capital-efficient leveraged investing using collateralised margining (something inherent with futures markets), the manager with the highest risk-adjusted return should be in hot demand and why as a consequence, many CTAs and hedge fund managers use the Sharpe ratio as their long-term performance objective function.

[9] M/E is usually defined as the average amount of margin required by the FCM (M) divided by the un-notionalised management account size of the CTA (P), thus, average $M/E\% = (M \times 100)/P$.

5.2 ABSOLUTE RISK-ADJUSTED RETURN METRICS

In this section, we will describe the group of absolute metrics currently used within the industry and use an example to show how each can be calculated for 10 hypothetical hedge funds (see Box 5.1) and to present a conclusion on the final results and the differences between methods.

Box 5.1 10 Hypothetical Hedge Funds

The table below shows 10 hypothetical hedge monthly returns between 2005 and 2010* classified as follows:

- Commodity Trading Advisors: CTA1, CTA2 and CTA3.
- Global Macro funds: GM1 and GM2.
- Long-Short funds: LS1, LS2 and LS3.
- Market Neutral funds: MN1 and MN2.

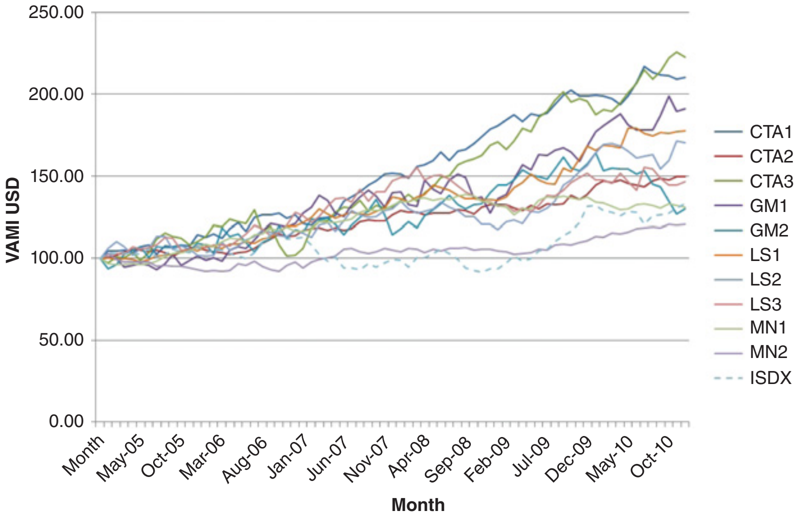
	A	B	C	D	E	F	G	H	I	J	K	L
1		CTA1	CTA2	CTA3	GM1	GM2	LS1	LS2	LS3	MN1	MN2	Average
2	**M1**	12.82	6.86	14.21	11.81	5.52	9.84	9.57	6.91	4.65	3.32	**8.55**
3	**M2**	8.74	5.27	12.69	14.14	15.16	7.31	11.51	10.64	5.60	5.86	**9.69**
4	**M3**	0.07	0.35	-0.55	-0.02	0.13	0.89	0.11	0.23	0.10	0.11	**0.14**
5	**M4**	0.04	-0.85	1.71	-0.11	0.36	0.61	-0.57	-0.01	-0.29	0.23	**0.11**
6												
7		CTA			Global Macro		Long Short			Market Neutral		
8	Month	CTA1	CTA2	CTA3	GM1	GM2	LS1	LS2	LS3	MN1	MN2	
9	Jan-05	4.68	0.10	-2.69	0.89	-6.66	0.37	5.80	2.36	-3.09	2.81	
10	Feb-05	-0.18	1.79	3.61	-0.90	2.93	-1.06	4.27	0.32	-0.87	-3.51	
11	Mar-05	0.66	2.04	-0.28	-5.38	6.68	0.62	-3.11	-0.44	0.19	-1.56	
12	Apr-05	-1.45	1.34	2.95	1.33	2.26	-0.62	-4.67	4.89	0.56	-0.07	
13	May-05	3.11	-0.16	-4.83	0.88	2.21	-1.40	4.12	-1.54	-0.82	0.38	
14	Jun-05	1.10	-1.40	5.01	-1.02	-6.26	1.36	-0.31	1.02	0.97	-0.07	
15	Jul-05	-5.11	1.15	8.03	-2.88	6.66	2.02	6.87	1.12	1.50	-2.85	
16	Aug-05	4.91	-1.20	3.25	3.81	-0.72	0.74	0.31	4.32	2.41	0.62	
17	Sep-05	-0.70	-1.99	-1.96	5.93	0.38	0.54	-6.94	-0.10	2.23	-0.46	
18	Oct-05	1.33	2.20	-0.75	-6.20	1.30	1.77	1.97	-6.56	0.51	-0.33	
78	Oct-10	-0.31	-0.69	3.89	6.19	-6.24	-0.45	3.36	-2.65	1.87	1.38	
79	Nov-10	-0.85	1.27	1.42	-4.46	-5.71	0.78	7.45	-0.01	-0.74	-0.14	
80	Dec-10	0.45	0.10	-1.33	0.59	2.48	0.10	-0.54	1.27	-0.75	0.50	

10 hypothetical hedge funds and their first four moments

CTA3 is seen to be the best performer in terms of the first moment (M1) with a return of 14.21%, closely followed by CTA1 at 12.82%. The clear poorest performer based on return is MN2 with a value of 3.32%.

The 10 hypothetical hedge funds will be used for the remainder of the chapter in order to demonstrate how the various metrics and asset pricing models can be used to investigate hedge fund performance. The results will also highlight the key differences in each approach as each technique will produce different rankings for each fund.

The VAMIs for each hypothetical hedge fund are shown below:



VAMIs of the 10 Hypothetical Hedge Funds (shows the *arithmetic* progression over time of a 100 USD starting investment in each fund)

*The monthly returns were simulated using Geometric Brownian Motion (GBM) and resemble a typical range of values found with such hedge fund strategies. The advantage of this approach is that you get a full data set for learning and experimentation purposes which can then be replaced by real data sets of your own choosing.

Most risk-adjusted return metrics may be generalised using the following formulation:

$$\text{Risk-Adjusted Return Metric} = \text{Return proxy/Risk proxy} \qquad (5.3)$$

Usually the return proxy is the estimated ex-post annualised return for the hedge fund less the annualised risk-free rate or some minimum acceptable rate. The risk proxy the most widely used is usually the volatility of returns since many investors are comfortable with viewing a manager in mu-sigma space. However, it has drawbacks due to its simplicity, for example higher moments such as kurtosis are neglected, the distribution could be asymmetrical (skewed) and drawdowns could turn out to be much worse than assumed. That is, volatility is essentially a measure of normal or Gaussian risk in an abnormal non-Gaussian real world of risk. As a result, the risk measure shown in the denominator in Equation 5.3 has been changed and re-defined by various researchers over the years to encompass more complex and meaningful measures as demanded by

```
%File: SR.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xls','10 Hedge Funds');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xls','15 Factors');

rfaverage = mean(rfreturn(:,6)); %average risk-free rate
sr = sharpe(returns,rfaverage/12)*sqrt(12); %Sharpe ratio

%sort data
sr = num2cell(sr);
data = [names(1,2:11)', sr']; %join columns
data = sortrows(data,2); %sort rows by Sharpe ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.1 MATLAB® code for the calculation of the Sharpe ratio

sophisticated investors. These metrics will be described and modelled in the following sub-sections.

5.2.1 The Sharpe Ratio

The Sharpe ratio[10] is defined by Sharpe (1994) as:

$$Sharpe = \frac{R_P - R_F}{\sigma_P} \tag{5.4}$$

where R_P is the annualised return, R_F is the annualised risk-free rate (e.g. using the T-bill as a proxy) and σ_P is the volatility of the returns and as such the Sharpe ratio introduces the concept of a benchmark to the numerator.

In other words, the investor wants to earn at least the risk-free rate and since they want the highest Sharpe ratio, this will begin to penalise hedge funds whose returns are low. Also, any fund with annualised returns less than the risk-free rate will have a negative Sharpe ratio regardless of the fund volatility.

The results are shown in Table 5.2.

As can be seen the Sharpe ratio produces a ranked list which captures very well in one metric the concept of the risk-adjusted return or return per unit risk. This metric is already a vast improvement on using just the annualised returns for performance ranking since the hedge fund manager can use leverage (as opposed to traditional fund managers who cannot use leverage), and so they need to be ranked relative to their peers in the risk-adjusted sense. As mentioned earlier, a bad manager can simply use leverage to gain superior results but when

[10] William Sharpe (1934–) is mostly known for the development of the Capital Asset Pricing Model (CAPM).

Table 5.2 Sharpe ratio for the 10 hypothetical hedge funds output

CTA1	0.98
CTA3	0.78
LS1	0.76
GM1	0.53
CTA2	0.49
LS2	0.46
LS3	0.25
GM2	0.08
MN1	0.06
MN2	−0.17

the SR is measured, their results would not seem superior. It is for this reason primarily that the SR is a very commonly used statistic in the world of hedge funds – allowing a manager or investor to be able to get the picture quickly of whether a strategy or fund is worth pursuing. In a sense, the SR is the first port of call in any hedge fund analysis and if an SR looks high, further analysis would then be undertaken to assess other statistical characteristics of the fund.

The disadvantage of the SR is that it is absolute in the sense it will not measure the manager meaningfully against a market vector benchmark, just the risk-free rate. However, other measures will be described in the following sections which are relative versus a market benchmark for ranking purposes. Other disadvantages are that the SR is based exclusively on mean-variance, and so higher moments and co-moments are neglected and are assumed to be zero. The SR therefore assumes that the returns are Gaussian which is not always the case for hedge funds, especially the more exotic relative value strategies as analysed by Lhabitant (2007).

5.2.2 The Modified Sharpe Ratio

The Modified Sharpe ratio (MSR) as defined by Gregoriou and Gueyie (2003) is defined as:

$$MSR = \frac{R_P - R_F}{|MVaR_p|} \tag{5.5}$$

The MSR takes the Sharpe ratio one stage further in complexity since it takes into account the higher moments of the distribution, i.e. skewness (S) and excess kurtosis (K). Since the Sharpe ratio is based only on return-volatility space, the MSR was developed to capture higher moments in the risk measure of the denominator using the Cornish-Fisher expansion. The Cornish-Fisher expansion is given by:

$$\tilde{z}_\alpha = z_\alpha + \frac{1}{6}\left(z_\alpha^2 - 1\right)S + \frac{1}{24}\left(z_\alpha^3 - 3z_\alpha\right)(K-3) - \frac{1}{36}\left(2z_\alpha^3 - 5z_\alpha\right)S^2 \tag{5.6}$$

where the Modified VaR ($MVAR_p$) measure is given by:[11]

$$MVaR_P(1-\alpha) = abs\left(-\left(\mu + \tilde{z}_\alpha \sigma\right)\right) \tag{5.7}$$

where z_α is the z-statistic corresponding to a one-sided standard normal PDF for $1-\alpha$ confidence (e.g. z = 2.33 for $1-\alpha = 0.01$, i.e. 99% confidence). For the Gaussian case $S = 0$ and $K = 0$ so $\tilde{z}_\alpha = z_\alpha$.

Source 5.2 shows the MATLAB® source code to calculate the MSR.

Table 5.3 shows that results for the ranked MSR when applied to the 10 hypothetical hedge funds versus the M1/M2 and Sharpe ratios.

Interestingly CTA1, LS1 and CTA3 dominate the top three rankings but not in the same order as for the SR – showing that there is extra information in the higher moments of the MSR which has led to an improved ranking beyond the mean-variance paradigm. It is also seen that the MSR will rank funds differently depending on the level used for the VaR computation.

5.2.3 The Maximum Drawdown Ratio

The Maximum Drawdown ratio (DD ratio) is given by:

$$DD\ Ratio = \frac{R_P - R_F}{|\max\ DD|} \tag{5.8}$$

The DD ratio is yet another enhanced variant of the Sharpe ratio but this time the measure of risk in the denominator is determined by the absolute value of the maximum historical drawdown. The maximum drawdown is defined as the maximum amount lost in either VAMI (or NAV) terms from the highest preceding high to the lowest low during the period that the hedge fund remains under water (it has not recuperated its value to that above the last highest high – the high water mark). Once the funds value has recuperated its losses and has achieved a net new high, then it is said to be out of its drawdown. The duration of this maximum drawdown period is called the maximum drawdown duration and the difference between the highest high VAMI to the lowest low VAMI during this period is the magnitude of the maximum drawdown. Note that the DD ratio uses the absolute value of the maximum drawdown so that it is always a positive number since it is usually reported as a negative value (i.e. a loss).

[11] MVaR will be discussed in further detail in Chapter 7.

```
%File: MSR.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xlsx','15 Factors');

rfreturn = rfreturn(:,6); %get the risk-free rate
rfaverage = mean(rfreturn); %mean risk-free rate

alpha = input('Input VaR probability e.g. 0.05 for 5%:');
%input VaR one-sided alpha probability

m1 = mean(returns);  %mean
m2 = std(returns);  %standard deviation
m3 = skewness(returns);  %skewness
m4 = kurtosis(returns);  %kurtosis

z = norminv(alpha,0,1);  %one-sided z-value for the VaR alpha
zcf = z+(1/6)*((z^2)-1)*m3+(1/24)*((z^3)-3*z)*m4-(1/36)*(2*(z^3)-
5*z)*(m3.*m3); %modified Cornish Fisher z-value

mv = abs(m1+zcf.*m2)*sqrt(12);  %absolute values of modi-
fied VaR annualised
msr = (m1*12-rfaverage)./mv;  %modified Sharpe ratio

%sort data
msr = num2cell(msr);
data = [names(1,2:11)', msr']; %join columns
data = sortrows(data,2); %sort rows by Information ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.2 MATLAB® code for the calculation of the Modified Sharpe ratio

Table 5.3 Ranked results for MSR for Modified VaR using 95% and 99% levels

CTA1	0.85	LS1	0.41
LS1	0.81	CTA1	0.38
CTA3	0.56	CTA3	0.23
CTA2	0.43	CTA2	0.22
GM1	0.39	GM1	0.19
LS2	0.34	LS2	0.18
LS3	0.18	LS3	0.09
GM2	0.06	GM2	0.03
MN1	0.05	MN1	0.02
MN2	−0.12	MN2	−0.06

```
%File: DD.m
%import XL data
[returns,names] = xlsRead('hfma_matlab_data.xlsx','10 Hedge Funds');
[rfreturn,rfname] = xlsRead('hfma_matlab_data.xlsx','15 Factors');

rfreturn = rfreturn(:,6); %get risk-free rates
rfaverage = mean(rfreturn); %mean risk-free rate
dd = maxdrawdown(fVAI(returns))*100; %max. drawdown

for i=1:10;
    dd(i) = mean(returns(i)*12 -rfaverage)/dd(i); %drawdown ratio
end;

%sort data
dr = num2cell(dd);
data = [names(1,2:11)', dr']; %join columns
data = sortrows(data,2); %sort rows by drawdown ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.3 MATLAB® code for calculating maximum drawdown

A user-defined VBA function to calculate the maximum absolute drawdown is shown in Source 5.3. Note that the input range is not monthly returns in this case but the VAMIs constructed from the monthly returns.

In Source 5.3 we use the maximum historical drawdown whereas other variants such as the Sterling ratio use an average of the most significant drawdowns, and as such will not penalise a manager with one major drawdown and some smaller ones compared to a manager with several bad drawdowns. The Burke ratio uses the square root of the sum of the squares of each drawdown, again with the idea in mind of penalising more significant extended drawdowns compared with several milder ones. The ranking results are shown in Table 5.4.

Table 5.4 The DD ratio output

CTA1	9.25
LS3	2.92
LS1	1.36
GM2	1.32
MN2	1.06
CTA3	0.16
GM1	−1.32
MN1	−1.39
CTA2	−1.55
LS2	−4.96

Table 5.5 The absolute group of metrics and their relative rankings

RANK	SR		MSR 0.05		MSR 0.01		DD	
1	CTA1	0.98	CTA1	0.85	LS1	0.41	CTA1	9.25
2	CTA3	0.78	LS1	0.81	CTA1	0.38	LS3	2.92
3	LS1	0.76	CTA3	0.56	CTA3	0.23	LS1	1.36
4	GM1	0.53	CTA2	0.43	CTA2	0.22	GM2	1.32
5	CTA2	0.49	GM1	0.39	GM1	0.19	MN2	1.06
6	LS2	0.46	LS2	0.34	LS2	0.18	CTA3	0.16
7	LS3	0.25	LS3	0.18	LS3	0.09	GMl	−1.32
8	GM2	0.08	GM2	0.06	GM2	0.03	MN1	−1.39
9	MN1	0.06	MNl	0.05	MN1	0.02	CTA2	−1.55
10	MN2	−0.17	MN2	−0.12	MN2	−0.06	LS2	−4.96

As seen from Table 5.4, there is another change of placing in the top three funds, and while CTA1 retains its place at the top of the ranking, LS3 is now second choice and LS1 third according to the DD ratio rankings.

5.3 MARKET MODEL RISK-ADJUSTED RETURN METRICS

In this section, we will consider metrics based around the Market Model. The Market Model differs from the CAPM in the sense that it is based around a time series regression of excess returns of individual assets on the excess returns of some reference market benchmark or aggregate index which is written as:

$$R_i = \alpha_i + R_F + \beta_i(R_M - R_F) + \varepsilon_i \tag{5.9}$$

or expressed in terms of risk preference:

$$R_i - R_F = \alpha_i + R_F + \beta_i(R_M - R_F) + \varepsilon_i \tag{5.10}$$

According to the CAPM all fairly priced assets with respect to publicly known information should then fall on the Security Market Line (SML) which is plotted in Figure 5.3.

From Figure 5.3 manager B is seen to exhibit greater than expected performance (positive alpha) and manager A is seen to exhibit less than expected performance (negative alpha) with respect to the market benchmark.

From the CAPM we can estimate this measure of a hedge fund's added value since:

$$\alpha_P = R_P - E^{CAPM}(R_M) \tag{5.11}$$

$$= (R_P - R_F) - \beta_P(R_M - R_F) \tag{5.12}$$

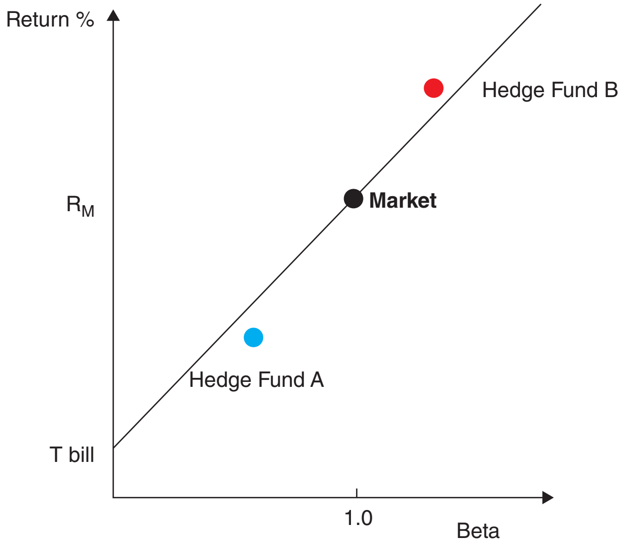


Figure 5.3 The SML for the market (zero alpha), Hedge Fund A and Hedge Fund B

where $E^{CAPM}(R_M)$ is the "unconditional expectation operator" (the expected return on the market, not conditioned on anything else occurring) applied to the benchmark or market index (CAIA Level 1 Study Handbook (2009)).

Equation 5.12 allows the calculation of a manager's added value relative to some beta controlled quantity of systematic market return premium. Equation 5.13 is commonly referred to as Jensen's alpha:

$$\alpha_J = (R_P - R_F) - \beta_P(R_M - R_F) \tag{5.13}$$

5.3.1 The Information Ratio

The information ratio as described by Goodwin (1998) is similar to the Sharpe ratio in that it is based on a fund's return and volatility. However, where the Sharpe ratio uses a risk-free rate as a benchmark for the numerator, the Information ratio goes one step further and uses a market reference benchmark vector. If R_{Pt} is the return on a fund in period t and R_{Bt} is the return on a benchmark portfolio in period t then Δ_t the excess return can be written as:

$$\Delta_t = R_{Pt} - R_{Bt} \tag{5.14}$$

and $\bar{\Delta}$ is the arithmetic average of excess returns from $t = 1$ to T

$$\bar{\Delta} = \frac{1}{T}\sum_{t=1}^{T}\Delta_t \tag{5.15}$$

```
%File: IR.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xlsx','15 Factors');

rfreturn = rfreturn(:,6); %get risk free-rates
rfaverage = mean(rfreturn); %mean risk-free rate

ir = inforatio(returns,rfreturn/12)*sqrt(12); %information ratio

%sort data
ir = num2cell(ir);
data = [names(1,2:11)', ir']; %join columns
data = sortrows(data,2); %sort rows by information ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.4 MATLAB® code for calculating the information ratio

where σ_Δ is the standard deviation of the excess returns from the benchmark (or tracking error[12]) written as:

$$\sigma_\Delta = TE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\Delta_t - \bar{\Delta})^2} \tag{5.16}$$

The Information ratio is then given by:

$$Information = \frac{\bar{\Delta}}{TE} \tag{5.17}$$

Table 5.6 shows the information ratio using a hypothetical global equity benchmark:

Table 5.6 The information ratio output

CTA1	0.98
CTA2	0.48
CTA3	0.78
GM1	0.53
GM2	0.08
LS1	0.75
LS2	0.46
LS3	0.25
MN1	0.06
MN2	−0.16

[12] This is also known as the tracking error which was introduced in Chapter 2 (see Section 2.4.1).

Table 5.6 shows that the rankings change once again. CTA1 is still the highest ranked and MN2 still the lowest ranked. However, CTA2 is in second place with CTA3 in third. In other words, some funds are starting to look better than before in the presence of a benchmark and some are starting to slip down the rankings.

5.3.2 The Treynor Ratio

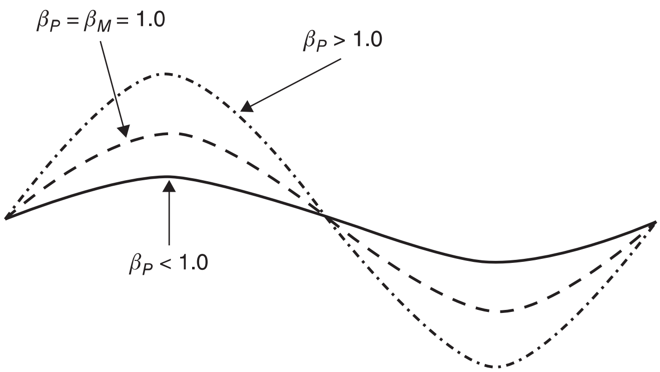
The Treynor ratio as introduced by Treynor (1962) is again of the same form as the generalised risk-adjusted return given by Equation 5.3. However this time risk is defined as the beta of the fund relative to a benchmark, such that:

$$\textit{Treynor Ratio} = \frac{\alpha_P}{\beta_P} = \frac{R_P - R_F}{\beta_P} \tag{5.18}$$

The beta of the portfolio β_P can be calculated two ways. The first method uses Ordinary Least Squares (OLS) to estimate the slope of the regression line (method two being Equation 5.19) as already discussed in Chapter 3 (see Section 3.9). Box 5.2 shows a detailed explanation of beta and its estimation.

Box 5.2 Beta and Correlation

Beta is a very common measure for estimating the extent to which the hedge fund excess returns are influenced by that of the benchmark or market excess returns. The beta of the portfolio is the gradient (*m*) of the OLS regression line where the market excess returns are plotted on the *x*-axis and the hedge fund excess returns on the *y*-axis. The fund's beta can be visualised as a measure of its relative directional exposure to the benchmark, i.e. a kind of relative elasticity as shown in the diagram below:



How to visualise beta

When the hedge fund's beta is greater than 1.0, it has tendency to exhibit greater moves in its returns (both positive and negative) than those of the benchmark. When the fund's beta is equal to 1.0, the fund is the market held long (if the market gains 1%, the fund gains 1% and for all intents and purposes the fund is exactly matched to the market, i.e. it is the market). When the fund's beta is between zero and 1.0, the fund's returns are less than those of the market (this is the region of beta usually occupied by long-short funds). If the fund's beta is zero, the fund is technically market neutral and so whatever the benchmark does has no impact on the returns of the fund. If the fund's beta is between zero and −1.0, then the fund's returns have a tendency to be opposite to those of the fund, but in a less volatile fashion. If the fund's beta is −1.0, the fund is the benchmark held short and as such the fund models the benchmark exactly but in the opposite way (if the benchmark gains 1%, the fund loses 1% and vice-versa).

Depending on your objectives, you may have different groups of betas to analyse. For example, if you are analysing a group of market neutral managers versus a market benchmark you may well have a statistical distribution of beta values around zero. However, if you are a fund of fund manager analysing the betas of various hedge funds within a diversified fund of hedge funds portfolio, you will have a group of betas distributed around one. Either way, the beta is the *sensitivity* of the fund or asset relative to some systematic factor such as the market or collective index or factor as seen in the Merton I model (1990) of Equations 5.19 and 5.20:

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2} \tag{5.19}$$

$$\alpha_i = \beta_i(R_M - R_F) \tag{5.20}$$

Beta is however directly proportional to leverage which opens up a new set of problems for risk-adjusted return analysis when using beta since any equation with beta in it will be leverage dependent – a case which is usually undesirable if the manager is to be measured for his skill and not leverage.

The beta however can be re-scaled to create an unleveraged cash equivalent *correlation coefficient*:

$$\rho_i = \frac{\sigma_M}{\sigma_P}\beta_i \tag{5.21}$$

And so, in the world of hedge funds we see the emergence of the correlation coefficient as representing the sensitivity of a fund to a certain factor which has been risk-adjusted. The correlation coefficient value if multiplied by 100 will then show the percentage of the dependent variable accounted for by the source of returns or premium for the factor in question, e.g. the systematic returns associated with the market portfolio if the CAPM is used. The correlation squared or r-squared when multiplied by 100 measures the proportion of variance captured in the dependent variable by the factor returns.

The results for the Treynor ratio are shown in Table 5.7.

The Treynor ratio is a metric more commonly used for active equity portfolios (like the Information ratio). This is due to the fact that if we are to use a benchmark for relative performance measurement, then the benchmark must be relevant. Since we are using a global equity market benchmark, it follows that the CTAs and global macro funds have little or no intuitive exposure to the benchmark and so to use it for any of the risk-adjusted metrics

```
%File: TR.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');
[market,marketname] = xlsread('hfma_matlab_data.xlsx','15 Factors');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xlsx','15 Factors');

market = market(:,1); %get PSDX as market return
benchmark
meanfund = mean(returns)*12; %mean fund return
meanmarket = mean(market)*12; %mean market return
rfreturn = rfreturn(:,6); %get risk-free rates
rfaverage = mean(rfreturn); %mean risk-free rate

for i=1:10;
    VCV = cov(returns(:,i),market); %VCV matrix
    fundvol = sqrt(VCV(1,1))*sqrt(12); %fund volatility
    marketvol = sqrt(VCV(2,2))*sqrt(12); %market volatility
    gearing = fundvol/marketvol; %gearing vs. market
    beta = VCV(1,2)/VCV(2,2); %fund beta
    cashbeta = beta*gearing; %cash beta
    tr(i) = (meanfund(i)-rfaverage)/beta; %Treynor ratio
end

%sort data
tr = num2cell(tr);
data = [names(1,2:11)', tr']; %join columns
data = sortrows(data,2); %sort rows by Treynor ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.5 MATLAB® code for calculating the Treynor ratio

Table 5.7 Treynor ratio output

LS2	62.15
CTA1	60.69
GMl	50.73
LSl	44.44
CTA2	32.60
MN2	10.40
MNl	−11.77
CTA3	−28.63
GM2	−36.47
LS3	−61.51

above would be practically meaningless. However, since most equity markets are highly correlated, the global benchmark would be efficient at capturing the returns associated with a passive investment in this asset class. As such, it may be a suitable metric for the equity hedge funds LS1, LS2, LS3 and the market neutral hedge funds MN1 and MN2 all of which are assumed to be global equity hedge funds.

The outcome is that we are starting to see the limitations of using risk-adjusted returns with benchmarks for general hedge fund performance measurement. From Table 5.8 we can see that LS2 is now ranked the highest and LS3 the bottom ranked hedge fund. This is primarily because LS2 has a beta of 0.08 and LS3 has a beta of −0.04. In fact, the results can be highly misleading, since both betas may be no different from zero. In such a market neutral case, the Treynor ratio approaches infinity. The higher the maxima and minima of the ranks of the Treynor ratio therefore, the more market neutral the fund is and so the manager's returns are therefore less influenced by the market which is a sign of outperformance with respect to the market. All in all, the Treynor ratio is difficult to interpret and clearly unsuitable for analysing non-equity hedge funds.

Table 5.8 Jensen's alpha ranking output

CTA3	10.18
CTAl	8.37
GMl	7.35
LSl	5.41
LS2	5.18
LS3	2.66
CTA2	2.48
GM2	1.16
MNl	0.38
MN2	−0.88

Box 5.3 Hypothesis Testing for the Significance of Beta

As seen above, LS2 has an estimated beta of 0.08. The question is whether this value of beta is significantly different from zero. In other words can we really believe the level of beta estimated at a 95% confidence? To find out, we need to perform a hypothesis test. The common approach to testing the strength of a beta estimate is by using the t-test statistic as discussed in Chapter 4 (see Box 3.3 in Chapter 3) and given by:

$$t = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}} \tag{5.22}$$

Since two parameters were estimated in the regression (the intercept and the gradient), the t-statistic has two degrees of freedom. The denominator $s_{\hat{b}_1}$ is the standard error. The hypothesis test (H-test) is usually carried out by determining a (critical or absolute) t value t_c which corresponds to the confidence interval of choice (e.g. typically 95% or 99%) in order that we can reject the null hypothesis if $t > t_c$. In our example we want to know if LS2 which has an estimated beta of 0.08 is significantly different from zero (i.e. the null hypothesis).

We test LS2 against the null hypothesis that its beta is no different from zero at 95% confidence, i.e. assuming that it is a market neutral fund with a beta of zero until proven that it is not with only a 5% higher chance of being wrong. Here we are testing the hypothesis that $H_0 \neq H_1$ and we therefore use the two-tailed t test. In Excel, for example, the standard error can be calculated using the following function:

$s_{\hat{b}_1}$ = **STEYX** (hedge fund excess returns, benchmark excess returns)

The critical two-tailed t-value is calculated using the following function:

t_c = **TINV** (probability, degrees of freedom)

Where probability refers to the probability of the significance level of the test, i.e. 0.05 for 95% and 0.01 for 99% confidence.

The degrees of freedom are the total number of observations minus two. So for our example, the degrees of freedom = 72 − 2 = 70. For the LS2 fund the standard error is 3.36 and critical t-value for 95% confidence and 70 degrees of freedom is 1.99 (t_c = 1.99). Since $b_1 = 0$, the test statistic t can be calculated as follows:

$$t = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}} = (0.08 - 0)/3.36 = 0.023$$

Since $t < t_c$ we cannot be sure at 95% if the fund has a beta which is not different from zero. In other words, we cannot reject the null hypothesis and therefore must assume the fund is market neutral even though it is classed as long-short which would suggest that it had a beta somewhere between zero and 1.0.

5.3.3 Jensen's Alpha

One of the key goals of quantitative manager selection is the computation of alpha. Alpha is the residual return left over once all known factors and risk adjustments have been accounted for and deducted from a fund's excess net return. As such, alpha is technically meant to measure the manager's absolute skill at using their *information set* to predict future price states in the various financial markets and instruments traded in their fund. Hedge fund managers are meant to be only active position takers since they are not paid to hold passive positions unless those passive positions are uniquely available to them only. It would be too easy otherwise and managers should not be remunerated for something you could easily do yourself.

$$\alpha_J = (R_P - R_F) - \beta_P(R_M - R_F) \tag{5.23}$$

Equations 5.13 and 5.23 are one and the same – Jensen's alpha. Although a basic measure, it is widely used in the world of traditional investing and is nevertheless a way to begin attacking the problem of alpha and beta separation (Jensen (1968)). The separation process helps break down the three components of return usually found within hedge funds, i.e. alpha (skill), the beta continuum as alluded to by Anson (2008) (skill to no skill) and the risk-free rate (no skill). Since managers of hedge funds usually charge high fees for their services (e.g. 2 and 20), it is of concern to investors that they have a good idea of what it is they are paying for.

Jensen's alpha is ranked and compared to other risk-adjusted metrics in Table 5.8.

Interestingly, CTA1 has been toppled from its premier position in the rankings by CTA3. In other words, CTA1 has been found guilty of *free-riding* on the market to a certain extent to produce a percentage of their returns. This is generally something frowned upon in the industry since a manager should not get paid for taking passive risk on an easily replicable asset class as the same exposure can usually be obtained cheaply using ETFs or futures positions held long and rolled as necessary. CTA3, which was third as ranked by Sharpe, is now in first place with a Jensen's alpha of 10.28%. While CTA1 is still in second place, GM1 keeps its place in third with a similar ranking to the Treynor ratio, but LS1 has been demoted to fourth place having been second in the Sharpe rankings. MN2 is still however the bottom ranked fund.

```
%File: JA.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');
[market,marketname] = xlsread('hfma_matlab_data.xlsx','15 Factors');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xlsx','15 Factors');

market = market(:,1); %get PSDX as market return
benchmark
meanfund = mean(returns)*12; %mean fund return
meanmarket = mean(market)*12; %mean market return
rfreturn = rfreturn(:,6); %get risk-free rates
rfaverage = mean(rfreturn); %mean risk-free rate

ja = portalpha(returns,market,rfreturn/12,'CAPM')*12; %Jensen's alpha

%sort data
ja = num2cell(ja);
data = [names(1,2:11)', ja']; %join columns
data = sortrows(data,2); %sort rows by Jensen's alpha
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.6 MATLAB® code for calculating Jensen's alpha

5.3.4 GH1 Metric

The drawbacks of Jensen's alpha are obvious in the sense that it does not facilitate the possibility of leverage used by the hedge fund manager – in other words it is directly dependent on leverage which is not a feature of a well specified model for hedge funds. The GH1 was developed by Graham and Harvey (1997) who were among the first academics to address the issues posed by risk-adjusted alpha. Along with the GH2 metric (described later in this chapter), they address the concept of either leveraging up or down the market benchmark to match the volatility of the hedge fund (as in the GH1 version) or leveraging up or down the hedge fund to match the volatility of the market benchmark (the GH2 version). We will look at the GH1 measure first since it is probably the less useful of the two measures when it comes to measuring the performance of hedge funds. The reason is that it de-leverages the market benchmark to match the volatility of the hedge fund. Since the hedge fund itself may be leveraged several times, this assumes feasible leveraging up or down of the market benchmark to some inflated or deflated arbitrary level of hedge fund leverage. The alpha is the difference between the returns of the fund and the leveraged market returns as seen in Figure 5.4 below. This metric is of interest depending on what it is you are looking for. The next set of metrics are

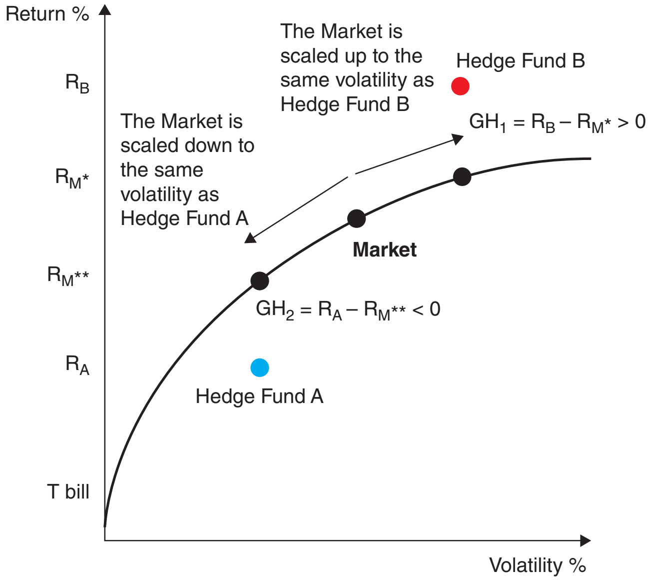


Figure 5.4 The GH1 metric and alpha measure

generally more useful since they create a cash-based market reference point for measurement of manager alpha and so are more attractive since they take leverage out of the equation and place the manager on a 1:1 leverage or gearing position relative to the market benchmark.

Of interest is the curvature of the efficient frontier between the market and the risk-free rate T-Bill due to the fact that the GH1 and GH2 measures take into account the curvature in the efficient frontier due to any covariance effects between the T-Bill and the market or any other funds (there is zero covariance only when the maturity of the cash instrument matches that of the in-sample estimation period).

Source 5.7 shows the code in MATLAB® to generate the ranked GH1 measure for the 10 hypothetical hedge funds in the Excel file `hfma_matlab_data.xlsx` – freely downloadable from darbyshirehampton.com along with many other data sets used in this book. The results are shown in Table 5.9.

5.3.5 The M2 Metric

The M-squared (or M2) was developed by Modigliani and Modigliani[13] as a risk-adjusted measure which could be more easily interpreted by the average

[13] See Modigliani (1997) and Modigliani and Modigliani (1997).

```
%File: GH1.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');
[market,marketname] = xlsread('hfma_matlab_data.xlsx','15 Factors');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xlsx','15 Factors');

market = market(:,1); %get PSDX as market return
benchmark
rfreturn = rfreturn(:,6); %get risk-free rates
rfaverage = mean(rfreturn); %mean risk-free rate

gh1 = portalpha(returns,market,rfreturn/12,'GH1')*12; %GH1 metric

%sort data
gh1 = num2cell(gh1);
data = [names(1,2:11)', gh1']; %join columns
data = sortrows(data,2); %sort rows by GH1 metric
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.7 MATLAB® code for calculating GH1

investor once the various hedge funds had been volatility adjusted for their average returns and then ranked. The intuition is that by using a benchmark with an estimated volatility the fund is leveraged or de-leveraged so that its volatility matches that of the benchmark. The leveraged or de-leveraged return of the fund is then reported and ranked. As such, the fund's return can be interpreted as the return that would have been produced had the fund's volatility been equivalent to that of the market benchmark. When various funds are processed in this way and ranked, it is easier for the less sophisticated investor to see how they outperform the benchmark to which they have had their risk profile matched.

Table 5.9 GH1 ranking output

CTA3	8.67
CTA1	7.67
GM1	6.13
LSl	4.83
LS2	4.15
CTA2	2.04
LS3	1.57
MNl	−0.20
GM2	−0.25
MN2	−1.55

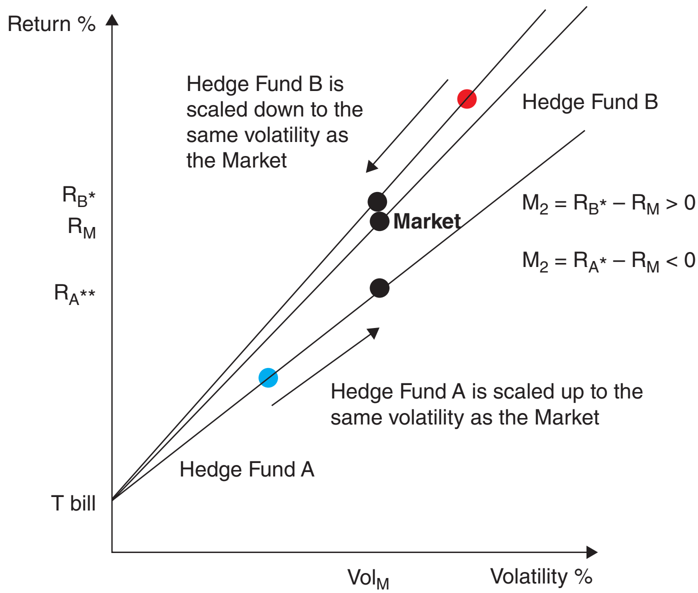


Figure 5.5 The M2 metric and alpha measure

For the 10 hypothetical hedge funds, the benchmark has an annualised return of 5.36% and a volatility of 10.9% for matching purposes. M2 is given by:

$$M^2 = \frac{\sigma_M}{\sigma_P}(R_P - R_F) - R_F \tag{5.24}$$

The ranked results are shown in Table 5.10.

It is seen that the rankings are identical to those of the Sharpe ratio. However, the main point of the M2 metric is that it is easy to interpret the rankings since they represent returns corresponding to volatilities that have been matched to

Table 5.10 M2 ranking output

CTA1	9.57
CTA3	7.45
LS1	7.20
GM1	4.73
CTA2	4.23
LS2	3.93
LS3	1.61
GM2	−0.19
MN1	−0.39
MN2	−2.88

```
%File: MM.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');
[market,marketname] = xlsread('hfma_matlab_data.xlsx','15 Factors');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xlsx','15 Factors');

market = market(:,1); %get PSDX as market return
benchmark
meanfund = mean(returns)*12; %mean fund return
meanmarket = mean(market)*12; %mean market return
rfreturn = rfreturn(:,6); %get risk-free rates
rfaverage = mean(rfreturn); %mean risk-free rate

mm = portalpha(returns,market,rfreturn/12,'MM')*12; %MM metric

%sort data
mm = num2cell(mm);
data = [names(1,2:11)', mm']; %join columns
data = sortrows(data,2); %sort rows by MM metric
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.8 MATLAB® code for calculating M2

that of the benchmark. Since the benchmark return is 5.36%, it follows that any fund with a return higher than 5.36% outperforms the market in the risk-adjusted sense, i.e. CTA1, LS1 and CTA3 (which is not instantly obvious from the Sharpe ratio rankings). Also, it is comforting for investors to know by how many percentage points their returns would be greater or less than the market for volatility matched hedge funds.

5.3.6 The GH2 Metric

As opposed to the GH1 metric, the GH2 metric is like the M2 metric in the sense that it leverages or de-leverages the hedge fund to the level of the cash market benchmark and in doing so gets rid of the effects of manager leverage. As such, the hedge fund can be directly compared to the market in the risk-adjusted sense and so any GH2 alpha as seen in the code below will be the risk-adjusted fund return minus any systematic market risk exposure as seen in Figure 5.6.

Table 5.11 shows the rankings of the 10 hypothetical hedge funds based on the GH2 metric.

Table 5.12 shows the combined results of the five metrics analysed and the respective rankings of the 10 hypothetical hedge funds.

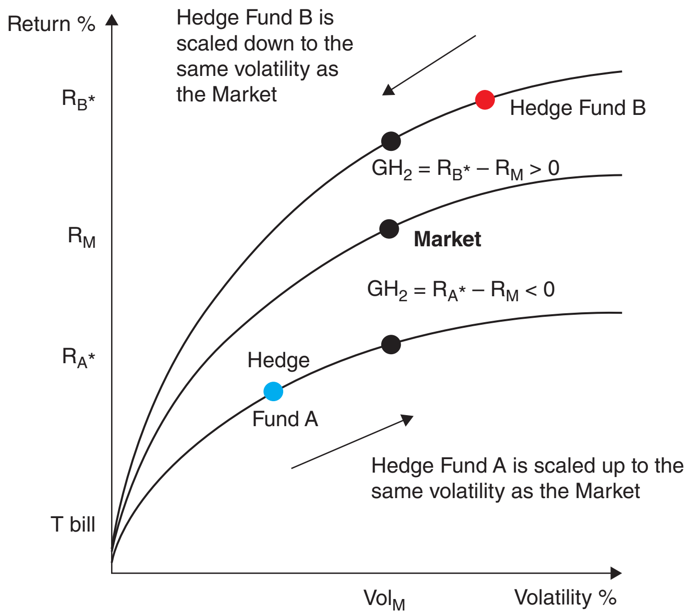


Figure 5.6 The GH2 metric and alpha measure

```
%File: GH2.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');
[market,marketname] = xlsread('hfma_matlab_data.xlsx','15 Factors');
[rfreturn,rfname] = xlsread('hfma_matlab_data.xlsx','15 Factors');

market = market(:,1); %get PSDX as market return benchmark
rfreturn = rfreturn(:,6); %get risk-free rates
rfaverage = mean(rfreturn); %mean risk-free rate

gh2 = portalpha(returns,market,rfreturn/12,'GH2')*12; %GH2 metric

%sort data
gh2 = num2cell(gh2);
data = [names(1,2:11)', gh2']; %join columns
data = sortrows(data,2); %sort rows by GH2 metric
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.9 MATLAB® code for calculating GH2

Table 5.11 GH2 output

CTA1	9.59
CTA3	7.45
LS1	7.17
GM1	4.73
CTA2	4.21
LS2	3.94
LS3	1.61
GM2	−0.19
MN1	−0.39
MN2	−2.87

Table 5.12 The market model group rankings

RANK	IR		TR		JA		GH1		MM		GH2	
1	CTA1	0.98	LS2	62.15	CTA3	10.28	CTA3	8.67	CTA1	9.57	CTA1	9.59
2	CTA2	0.48	CTA1	60.69	CTA1	8.37	CTA1	7.67	CTA3	7.45	CTA3	7.45
3	CTA3	0.78	GM1	50.73	GM1	7.35	GM1	6.13	LS1	7.20	LS1	7.17
4	GM1	0.53	LS1	44.44	LS1	5.41	LS1	4.83	GM1	4.73	GM1	4.73
5	GM2	0.08	CTA2	32.60	LS2	5.18	LS2	4.15	CTA2	4.23	CTA2	4.21
6	LS1	0.75	MN2	10.40	LS3	2.66	CTA2	2.04	LS2	3.93	LS2	3.94
7	LS2	0.46	MN1	−11.77	CTA2	2.48	LS3	1.57	LS3	1.61	LS3	1.61
8	LS3	0.25	CTA3	−28.63	GM2	1.26	MN1	−0.20	GM2	−0.19	GM2	−0.19
9	MN1	0.06	GM2	−36.47	MN1	0.38	GM2	−0.25	MN1	−0.39	MN1	−0.39
10	MN2	−0.16	LS3	−61.51	MN2	−0.88	MN2	−1.55	MN2	−2.88	MN2	−2.87

5.4 MAR AND LPM METRICS

This last group of metrics is based around the concept of separating out upside “good volatility” and the downside “bad volatility”. The variance and the standard deviation are two-sided measures of the degree of the spread in an asset’s returns. However, often it is desirable to be able to specify a Minimum Acceptable Return (MAR) threshold for a certain statistic in order to define some kind of minimum acceptable loss threshold. This last set of metrics consisting of the Sortino, Omega and Upside Potential ratios all use the concept of segregating desirable upside volatility and undesirable downside volatility with reference to the MAR (usually in percent) threshold level.

5.4.1 The Sortino Ratio

The Sortino[14] ratio is defined as:

$$Sortino = \frac{R_P - MAR}{\sqrt{\frac{1}{T}\sum\limits_{\substack{t=0 \\ R_P<MAR}}^{T} (R_{P,t} - MAR)^2}} \tag{5.25}$$

[14] Initially defined by Brian M. Rom in 1986 but later developed in the 1990s by Frank Sortino.

The Sortino ratio is another enhanced variant of the Sharpe ratio. The problem with the Sharpe ratio is that it uses volatility as a measure of risk for the entire time series of monthly returns. As a result, both large upswings and large downswings in value are penalised as they translate into higher volatility and a lower Sharpe ratio. As can be seen from Equation 5.27, the denominator uses the concept of the Minimum Acceptable Return (MAR) to set a reference point for the measurement of returns. Basically the MAR is a fixed point, e.g. 3.0%, which sets a minimum acceptable rate of return for the investor and in doing so splits the returns into two categories, i.e. those returns greater than or equal to MAR (the upside semi-deviation) and those returns less than MAR (the downside semi-deviation). Since the idea behind the concept of risk-adjusted returns is to try to have the least risk possible and the investor does not want their upside risk to be penalised, the Sortino ratio uses a measure of standard deviation in the denominator which only uses the downside returns in the calculation. Thus, the higher the Sortino ratio, the better the manager is at controlling downside returns while not being penalised for upside returns. As such it is a good metric for assessing the risk-management capabilities of a hedge fund manager.

The MATLAB® code to calculate the Sortino ratio is shown in Source 5.10.

The ranked results are shown in Table 5.13.

Table 5.13 shows that there seems to be a clear contender emerging from the risk-adjusted rankings, namely CTA1, since it has been the top ranked for all four methods used so far. CTA3 is also favoured over LS1 due to the MSR result – showing that the high kurtosis of CTA3 must be associated with good kurtosis, that is, high returns in the upside semi-deviation. Whilst having a negative skew, the returns in the downside semi-deviation below the MAR are

```
%File: SOR.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');

MAR = input('Input Minimum Acceptable Return (MAR) e.g. 5 for 5%:');
%input Minimum Acceptable Return (MAR)

md = mean(returns)*12-MAR; %Sortino numerator
sd = sqrt(lpm(returns*12,MAR,2)); %Sortino denominator
sr = md./sd;  %Sortino ratio

%sort data
sr = num2cell(sr);
data = [names(1,2:11)', sr']; %join columns
data = sortrows(data,2); %sort rows by Sortino ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.10 MATLAB® code for calculating the Sortino ratio

Table 5.13 Ranked Sortino ratio with MARs of 0% and 5%

LS1	1.02	CTA1	0.46
CTA1	0.88	LS1	0.39
CTA2	0.85	CTA3	0.31
CTA3	0.53	GMl	0.22
LS2	0.43	LS2	0.18
MN1	0.43	CTA2	0.17
GM1	0.42	LS3	0.08
LS3	0.32	GM2	0.01
MN2	0.27	MN1	−0.03
GM2	0.16	MN2	−0.11
MAR = 0%		MAR = 5%	

milder than those associated with the downside semi-deviation in LS1 on a risk-adjusted basis. As we go down the Sortino ranking, we also notice that GM1 is to be preferred over CTA2. In a nutshell, it shows how the population estimates of skewness and kurtosis can be misleading since CTA2 has a higher skewness (0.35 vs. −0.02 for GM1) and a lower kurtosis (−0.85 vs. −0.11 for GM1); meanwhile the Sortino ratio ranks GM1 higher than CTA2.

5.4.2 The Omega Ratio

The Omega ratio as defined by Keating and Shadwick (2002) is a non-parametric method, since it does not rely upon calculating the moments of a distribution of returns for determining the ranking of each hedge fund versus a threshold level. Developed in the early 2000s it has gained popularity with practitioners due to its inclusiveness of the full distribution of the returns in describing the relative ranking of a fund. As such, it incorporates all first four moments (and theoretically beyond). As mentioned earlier, investors like high odd moments and low even moments and the risks borne of holding an asset prone to fourth moment risk especially are non-negligible. Its popularity also stems from the fact that it inherently describes the risk-reward properties of the return distribution and so can be easily interpreted. The ratio considers returns below and above a specific loss *threshold* return level. In doing so, it places a threshold return on the unit distribution of returns (the return PDF with an area of one underneath) which have been accumulated by summing (taking the integral) from the left hand side to form the CDF, $F(x)$. Instead of using the positive and negative infinity signs to signify the limits on either side of the integral, we use a for the downside limit and b for the upside limit. The Omega ratio is written as:

$$\Omega(L) = \frac{\int_L^b (1 - F(x)dx}{\int_a^L F(x)dx} \tag{5.26}$$

```
%File: OM.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');

MAR = input('Input Minimum Acceptable Return (MAR) e.g. 5 for 5%:');
%input Minimum Acceptable Return (MAR)

on = lpm(-returns*12,-MAR,1); %Omega numerator
od = lpm(returns*12,MAR,1); %Omega denominator
om = on./od;  %Omega ratio

%sort data
om = num2cell(om);
data = [names(1,2:11)', om']; %join columns
data = sortrows(data,2); %sort rows by Omega ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.11 MATLAB® code for calculating the Omega ratio

The returns below a certain threshold are considered as losses, and the returns above are considered profits. It follows from Equation 5.26 that a higher value of Omega for a given threshold is preferred over a lower value for a rational investor. As such when the Omega ratio takes a value of one, this represents the mean of the portfolio since it is the balancing point in terms of moments. The Omega ratio can therefore be used to create rankings based on various return thresholds. The rankings may change as a function of a perceived loss threshold which will vary from investor to investor depending on their appetite for risk which is interpreted here as the magnitude of return threshold. The Omega ratio is non-parametric and reflects all moments of the distribution so a fund with a high excess kurtosis will be ranked lower than one with no excess kurtosis assuming both funds have the same return, volatility and skewness.

Table 5.14 shows the ranked results for the Omega ratio for two different MAR values, 0% and 5%.

Table 5.14 Ranked Omega ratios for MAR levels of 0% and 5%

CTA1	2.94	CTA1	1.93
LS1	2.86	CTA3	1.70
CTA2	2.52	LS1	1.64
CTA3	2.23	GM1	1.41
MN1	1.81	LS2	1.32
LS2	1.81	CTA2	1.27
GM1	1.81	LS3	1.14
LS3	1.60	GM2	1.02
MN2	1.53	MN1	0.96
GM2	1.29	MN2	0.81
MAR = 0%		MAR = 5%.	

```
%File: UP.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');

MAR=input('Input Minimum Acceptable Return (MAR) e.g. 5 for 5%:');
%input Minimum Acceptable Return (MAR)

un=lpm(-returns*12,-MAR,1); %Upside Potential numerator
ud=sqrt(lpm(returns*12,MAR,2)); %Upside Potential denominator
up=un./ud;  %Upside Potential ratio

%sort data
up = num2cell(up);
data = [names(1,2:11)', up']; %join columns
data = sortrows(data,2); %sort rows by Upside Potential ratio
data = flipdim(data,1); %sort in descending order
display(data);
```

Source 5.12 MATLAB® code for calculating the Upside Potential ratio

5.4.3 The Upside Potential Ratio and Group Rankings

Sortino *et al*. (1999a, b) suggested the Upside Potential ratio as a hybrid metric which is somewhere between the Sortino ratio (same denominator) and the Omega ratio (same numerator) in design. The numerator can be thought of as the potential for success and so is termed the Upside Potential ratio.

From the MATLAB® Finance Toolbox User Guide, the lower partial moment function (lpm) in combination with the MAR procedure, the three ratios as discussed:

```
Omega = lpm( -Data, -MAR, 1)/lpm(Data, MAR, 1)
Sortino = (mean(Data) -MAR)/sqrt(lpm(Data, MAR, 2))
Upside = lpm(-Data, -MAR, 1)/sqrt(lpm(Data, MAR, 2))
```

Source 5.12 shows the MATLAB® code used to calculate the Upside Potential ratios with Table 5.16 showing the ranked comparison of the three metrics.

There exist several other techniques in the literature which have not been covered here, for example the Stutzer index (Stutzer 2000), the Sharpe-Omega ratio (Gupta, Kazemi and Schneeweiss 2003), AIRAP (Sharma 2004), Kappa (Kaplan and Knowles 2004) and Rachev ratio. Lo (2002) also reports a modified risk-adjusted ratio method adjusted for autocorrelation. All these methods are outside the scope of this book since their functions are not directly supported by MATLAB®. However, custom functions could be created using the MATLAB® programming language.

Table 5.15 Ranked Upside Potential ratios for MAR levels of 0% and 5%

LS1	1.57	LS1	0.99
CTA2	1.42	CTA1	0.96
CTA1	1.33	CTA2	0.79
LS2	0.97	CTA3	0.77
MN1	0.96	GM1	0.76
CTA3	0.95	LS2	0.75
GM1	0.93	LS3	0.66
LS3	0.87	GM2	0.58
MN2	0.78	MN1	0.57
GM2	0.71	MN2	0.47
MAR = 0%		MAR = 5%	

Table 5.16 Ranked Sortino, Omega and Upside Potential ratios for MAR levels of 0%

RANK		SO		OM		UP
1	LS1	1.02	CTA1	2.94	LS1	1.57
2	CTA1	0.88	LS1	2.86	CTA2	1.42
3	CTA2	0.85	CTA2	2.52	CTA1	1.33
4	CTA3	0.53	CTA3	2.23	LS2	0.97
5	LS2	0.43	MN1	1.81	MN1	0.96
6	MN1	0.43	LS2	1.81	CTA3	0.95
7	GM1	0.42	GM1	1.81	GM1	0.93
8	LS3	0.32	LS3	1.60	LS3	0.87
9	MN2	0.27	MN2	1.53	MN2	0.78
10	GM2	0.16	GM2	1.29	GM2	0.71
	MAR = 0%					

Table 5.17 Ranked Sortino, Omega and Upside Potential ratios for MAR levels of 5%

RANK		SO		OM		UP
1	CTA1	0.46	CTA1	1.93	LS1	0.99
2	LS1	0.39	CTA3	1.70	CTA1	0.96
3	CTA3	0.31	LS1	1.64	CTA2	0.79
4	GM1	0.22	GM1	1.41	CTA3	0.77
5	LS2	0.18	LS2	1.32	GM1	0.76
6	CTA2	0.17	CTA2	1.27	LS2	0.75
7	LS3	0.08	LS3	1.14	LS3	0.66
8	GM2	0.01	GM2	1.02	GM2	0.58
9	MNl	−0.03	MN1	0.96	MNl	0.57
10	MN2	−0.11	MN2	0.81	MN2	0.47
	MAR = 5%					

The metrics introduced are seen to be tools in a tool box and if used wisely are not a Pandora's box as described by Géhin[15] (2007). They offer the investor various ways of looking at either the absolute, relative out-performance versus a benchmark or a MAR for hedge funds. An investor who wants to set up a managed account with high notional funding might want to look at the DD ratio for example, since their concern may be focused on how much the CTA could lose from the outset of their investment. As such the investor seeks to minimise *margin call-at-risk*.[16] On the other hand a sophisticated pension fund investor may be more worried about the effects of higher moments and the asymmetry of the returns with respect to a minimum acceptable return. As such, their interests may be better served using the MSR or the Sortino ratio. Finally, it must be noted that these results only hold in an absolute environment, i.e. where the only benchmark is a static risk-free rate (which sometimes can be ignored and set equal to zero – especially if the investor is living in a period of economic stagnation where interest rates are very low, e.g. Japan during the 1990s and 2000s). As such, the results could turn out to be misleading especially if the managers have a high correlation to various benchmarks or indices.

5.5 MULTI-FACTOR ASSET PRICING EXTENSIONS

In Section 5.4 we introduced the M2 statistic which is given by:

$$M^2 = \frac{\sigma_M}{\sigma_P}(R_P - R_F) - R_F \tag{5.27}$$

The M2 approach as used in MATLAB® is probably the most intuitive and useful method for analysing the linear risk-adjusted sources of returns for hedge funds. This is primarily because the leverage multiplier:

$$L = \frac{\sigma_M}{\sigma_P} \tag{5.28}$$

acts as a de-leveraging multiplier effectively creating a cash equivalent un-leveraged hedge fund return. So long as the market benchmark is un-leveraged, the fund can be compared on a cash basis with the market. This removes any effects of amplification of returns derived from the use of leverage or gearing which the hedge fund manager may have used. Note also that the under-leveraged hedge fund manager is also punished, since their fund is multiplied up to that of the benchmark market in terms of leverage. In a nutshell, the leverage multiplier of Equation 5.30 acts in such a way as to set the hedge fund manager's volatility equal to that of the cash market benchmark allowing their

[15] *The Challenge of Hedge Fund Performance Measurement: A Toolbox Rather Than a Pandora's Box*, Walter Géhin, EDHEC-Risk Institute, January 2007.

[16] The probability of having margin calls on the managed account over a specified time, e.g. one year.

returns to be compared against those of the market factor on an equal footing cash basis.

From Source 5.8 above, the M2 alpha metric is calculated in MATLAB® using the following equation:

$$M^2 = \frac{\sigma_M}{\sigma_P}[(R_P - R_F) - \beta_P(R_M - R_F)] \tag{5.29}$$

Or expanding

$$M^2 = \frac{\sigma_M}{\sigma_P}(R_P - R_F) - \frac{\sigma_M}{\sigma_P}\beta_P(R_M - R_F) \tag{5.30}$$

It can be seen that the second term of the equation is the market risk premium multiplied by a term which is equivalent to the funds correlation coefficient with respect to the first factor or market factor:

$$\frac{\sigma_M}{\sigma_P}\beta_P = \rho_p \tag{5.31}$$

So the M2 alpha metric can be rewritten as follows:

$$M^2 = \frac{\sigma_M}{\sigma_P}(R_P - R_F) - \rho_P(R_M - R_F) \tag{5.32}$$

This statistic therefore is particularly suited for hedge fund performance analysis since it effectively creates a benchmark volatility adjusted portfolio return in the first term, and then creates a cash basis unleveraged benchmark market risk premium in the second term.

The upshot of Equation 5.32, is that the approach can be easily extended to accommodate further factor return premia which can be subtracted just as the first market factor premium was. In other words, a multi-factor version of the M2 statistic, with a further K market factors in addition to the market premium, would be:

$$M^2_{MF} = \frac{\sigma_M}{\sigma_P}(R_P - R_F) - \rho_P(R_M - R_F) - \rho_{F1}(R_{F1} - R_F) - \rho_{FK}(R_{FK} - R_F) \tag{5.33}$$

where:

$$\rho_{Fi} = \frac{\sigma_{Fi}}{\sigma_P}\beta_i \tag{5.34}$$

and

$$\beta_i = \frac{\sigma_{PFi}}{\sigma^2_{Fi}} \tag{5.35}$$

In the same way as in the Merton I model where:

$$\alpha_i = \beta_i(R_M - R_F) \tag{5.36}$$

and

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2} \tag{5.37}$$

It is seen that the various factors can be found using multiple linear regression and then each factor premium can be multiplied by their own de-leveraging factor as shown in Equation 5.33 to bring each factor beta to a cash equivalent beta, which is seen to be their correlation coefficient. This technique is a logical follow on to the treatment of the first market factor in the M2 alpha metric in MATLAB® although similar results will be obtained from the GH2 approach. Each additional factor other than the main market factor is then effectively leveraged or de-leveraged to the market volatility level, allowing all effects of leverage again to be neutralised. On a risk-adjusted basis, the managers' alpha is thus further reduced by the subtraction of extra de-leveraged factor exposures. The factor betas are generally derived in the usual way using multiple regression and checking r-squared values for hypothesis testing in order to accept or reject a factor.

In MATLAB®, the command `[b, bint, r, rint, stats]= regress(Y,X)` will produce a vector b which contains the *K* beta coefficients if *K* factors have been used in the X matrix. The Y matrix is the vector of the dependent variable values. The stats matrix contains values of r-squared, the F-statistic and its p-value for hypothesis testing.

5.5.1 The Choice of Factors

As documented in Section 5.5, there exist in theory *K* factors which may explain the performance of a hedge fund over and above the market's excess returns and the risk-free rate. A *benchmark* is a reference portfolio which is closely representative of the risk factor exposures of a fund. An *index*, however, is a reference portfolio that is representative of one or more risk factors or styles, e.g. a style index or a sector index. Indices make good benchmarks if they are therefore representative of the styles of the fund. A good set of benchmark qualities as proposed by Bailey (1990) include being:

- Unambiguous
- Investable
- Measurable
- Appropriate
- Reflective of current investment opinions
- Specified in advance.

Table 5.18 Subgroup 1: Beta factors

#	Abbreviation	Factor
1	PSDX	Passive Global Stock Index
2	S&P 500 Index	The S&P 500 US Equity Index
3	PBond DX	Passive Global Bond Index
4	PCom DX	Passive Long Global Commodity Index
5	PUSD DX	Passive Long US Dollar Index
6	Rf	Risk-Free Rate

Table 5.19 Subgroup 2: Industry reference alternative beta factors

#	Abbreviation	Factor
7	CTA Index	Commodity Trading Advisor Index
8	LS Index	Long-Short Equity Index

Tables 5.18, 5.19, 5.20 and 5.21 list the 15 hypothetical factors accompanying this book which have been further sub-divided into four subgroups (see *Preface*).

Subgroup 1 is indicative of global passive indices and as such is classed as a set of beta factors. Many indices exist in reality for which this set is a proxy and which are easily investable at low cost through ETFs and futures, e.g. Wiltshire, MSCI, Dow Jones, Barclays etc.

Subgroup 2 is representative of typical indices based on the reported returns from the various managers within each style group. This is a hypothetical proxy group representing two main styles, namely Long-Short equity hedge funds

Table 5.20 Subgroup 3: Fama-French-Carhart factors

#	Abbreviation	Factor
9	Val - Gr	Value minus Growth
10	SC - LC	Small Cap minus Large Cap
11	Mom	Momentum

Table 5.21 Subgroup 4: Active alternative beta factors

#	Abbreviation	Factor
12	ASDX	Active Global Stock Futures Index
13	ABDX	Active Global Bond Futures Index
14	ACDX	Active Global Commodity Futures Index
15	AFDX	Active Global Foreign Exchange Futures Index

and CTAs. These indices exist in investable (e.g. EDHEC Investable Hedge Fund Indices, Lyxor Investable Strategies etc.) and non-investable forms (e.g. Eurekahedge, IASG etc. – actual recorded past month and historical non-traded index reporting agencies). Since the indices we use are assumed to be investable and the exposures to underlying markets are by definition active in nature since they are composed of hedge funds, Subgroup 2 could be classed as sophisticated alternative beta.

Subgroup 3 shows the three classic factors introduced by Fama and French (1992) and later Carhart (1997). These factors differ from the other factors introduced since they are not directly investable. They remain a reasonably valid method for explaining the increased returns experienced by investing long value and short growth stocks, long small cap and short large cap (the difference reflects largely taking higher credit risks with the small cap companies hence it reflects additional risk premia earned from endogenous risk factors). Momentum measures the abnormal returns borne from holding long certain stocks within a stock market basket and shorting others depending on their recent performance.

Subgroup 4 is representative of alternative beta indices which are either available commercially or which could be replicated in-house using a transparent systematic trading approach, e.g. trend following using a 30-day moving average applied to futures markets. Four indices are proposed which cover the four main financial markets within the managed futures/CTA industry, i.e. stock index, bonds, commodities and foreign exchange.

- The factors in Subgroup 1 are beta and virtually costless to access via ETFs and futures markets.
- The factors in Subgroup 2 represent sophisticated alternative beta and represent the average industry performance for hedge funds depending on the style.
- The factors in Subgroup 3 are included for completion and remain relevant to mainly long-short performance analysis and do not represent true investable indices.
- The factors in Subgroup 4 are unsophisticated alternative beta and have been generated from a basic systematic trend following rule. They can be accessed via low costs funds or can be replicated in-house at low cost.

Amenc, Sfeir and Martellini (2003) document four types of factors models, namely:

1. *Implicit Factor Model* – Principal Component Analysis (PCA) is used to derive the inherent or implicit factors identified for the fund by calculating eigenvalues. This is the underlying approach to the EDHEC Pure Indices, i.e. taking a statistical extraction of the extent of each underlying hedge fund performance index to form a kind of consensus as to what must statistically

be the true or pure index, since it has the most components. PCA can be computed in MATLAB® using the pca command. The matrix of time series for analysis is as usual loaded into MATLAB®. Again, since leverage is assumed to be a variable for hedge funds, one way of standardising the input data is to use `zscore(X)` if necessary. The command is:

```
[coeff, score, latent, tsquared, explained] = pca(X)
```

`coeff` represents the PCA coefficents for the n-by-p data matrix X. Rows of X correspond to observations and columns to variables. The coefficient matrix is p-by-p. Each column of `coeff` contains coefficients for one principal component, where the columns are in descending order of component variance. See MATLAB® Statistics User Guide for further information or Lhabitant (2004).

2. Explicit *Macro Factor Model* – In this factor model, macroeconomic variables are either computed as predictive variables or used ex-post to gauge major market sensitivities to macroeconomic variables such as GDP expectations versus actual reported GDP (modelling the effect of the surprise), inflation etc.
3. *Explicit Micro Factor Model* – Microeconomic predictive variables are estimated and forecast, in much the same way the explicit macro factor model.
4. *Explicit Index Factor Model* – Each factor represents an explicit or observed value of some fund or index available as an ETF or future contract.

Other approaches in the literature of interest are the conditional approach of Kat and Miffre (2002), the long-short correlation risk-adjusted statistical arbitrage model of Hampton (2009, 2011) and the four moment extension of Favre and Ranaldo. Lhabitant (2004, 2007) provides key quantitative insights into the world of hedge funds as seen from the viewpoint of a fund of hedge funds manager. The EDHEC-Risk Institute provides interesting *research for business* articles and pure indices for most of the common hedge fund styles (Amenc and Martellini 2003) based on PCA available at www.edhec-risk.com.

6

Hedge Fund Classification

The principal idea behind classifying hedge funds is the ability to convert a set of hedge funds returns into groups so as to be able to visualise clear clustering and style group boundaries. For the fund of hedge fund (FoHF) manager this is indispensable – since clustering and classification are observable ways of scientifically studying hedge fund return data so that an empirical set of estimates can be produced for further hypothesis testing where necessary. As such, similarities emerge from the data – patterns from what would otherwise look very noisy to the naked eye. These patterns allow us to visually check a fund's grouping – to see how similar it is with respect to peers within a stated style. The main technique supported by MATLAB® is the *dendrogram* (Mantegna and Stanley (2000), Lhabitant (2004)). The dendrogram ("dendro" from the Greek meaning "tree") is the industry standard method of analysis of the natural hierarchy or taxonomy of a data set. This chapter allows us to apply the tools available in MATLAB® to delve deeper into the structure of hedge fund returns and gain a better understanding of their grouping and classification.

6.1 FINANCIAL INSTRUMENT BUILDING BLOCKS AND STYLE GROUPS

In the Eurekahedge database as of November 2011, there were 13,674 hedge funds – a huge number of funds considering that there are only eight main "building blocks" or instrument groups available for trading. Table 6.1 and Figure 6.1 show these building block instruments traded, the number of funds that primarily trade them and their total percentage of the total and the corresponding pie chart respectively.

It is seen that equity trading dominates the hedge fund universe with approximately 45% of all hedge funds engaged primarily in equity trading. Next is fixed income with 18%, which, if taken with equities, represents a total percentage of around 63% of all hedge fund trading. This is hardly surprising, since the capital markets for equity and debt are some of the deepest and most varied markets in the world. The remaining four instrument groups traded – currency, commodities, derivatives and cash – account for the remaining 35% with private equity and real estate being marginal instrument trading activities accounting for less than 2% of the total.

It is hardly surprising that many of these hedge funds will have similar characteristics since many of them will have to rely on the same set of building

Table 6.1 Instrument Traded, the Total Funds and % of Total

Instrument Traded	Total Funds	% of Total
Equities	6215	45.45
Fixed Income	2481	18.14
Currency	1118	8.18
Commodities	774	5.66
Derivatives	1611	11.78
Cash	1268	9.27
Private Equity	93	0.68
Real Estate	114	0.83

Source: Eurekahedge

blocks, e.g. diversified CTAs and Long-Short Equity managers. Eurekahedge also groups the hedge funds they follow into 16 "Style" groups or indices as seen in Table 6.2.

Figure 6.2 shows the evolution of the industry since 1991 with new funds being added regularly as seen on the right hand side, and funds being closed as seen in the left-hand part of the figure – a phenomenon which could be blamed primarily on the 2008 and beyond periods of excessive market turbulence, catching many "professional" managers unawares and leading to poor performance and the ensuing client exit and eventual closing of the funds in question. The graph nevertheless attests to the enormous evolutionary growth in the industry since 1991 – explaining the huge number of hedge funds in existence today.

6.2 HEDGE FUND CLUSTERS AND CLASSIFICATION

In the Eurekahedge database at the end of 2011 there were 13,674 "surviving" hedge funds split among 16 different styles but using only six main instruments

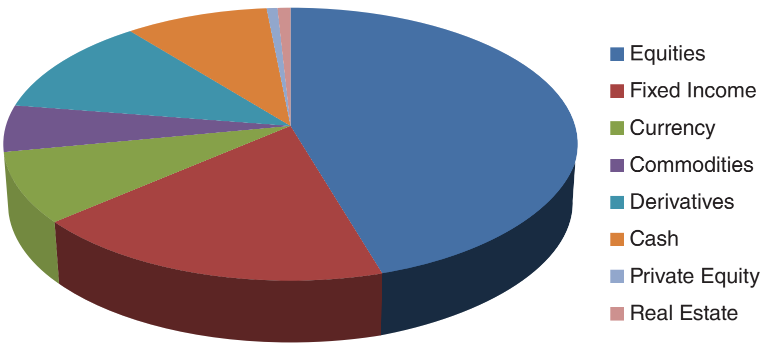


Figure 6.1 Pie chart of each main instrument group
Source: Eurekahedge

Table 6.2 The 16 Style groups and indices used by Eurekahedge

Eurekahedge Style Index	
1	Arbitrage
2	Bottom Up
3	CTA
4	Distressed Debt
5	Dual Approach
6	Event Driven
7	Fixed Income
8	Fund of Funds
9	Hedge Fund
10	Long-Short Equities
11	Long-Only Absolute
12	Macro
13	Multi-Strategy
14	Relative Value
15	Top-Down
16	Value

groups for trading (neglecting private equity and real estate which account for less than 2%). It can be a daunting prospect for the professional fund of hedge fund manager or investor to find themselves confronted with such a huge array of data if tasked with finding possible candidate hedge funds for analysis. As such, an approach widely taken in the industry is to first create empirical classification structures of the funds based on their reported monthly returns.

The most basic approach already covered in previous chapters is to create two dimensional return-volatility scatter plots. However, this approach tends to cluster managers with similar levels of gearing who produce similar rates

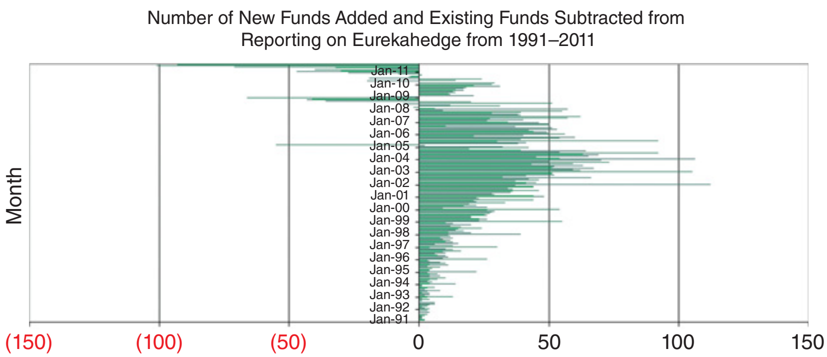


Figure 6.2 New funds added monthly (right side) and funds closed down (left side) from 1991 to 2011 in the Eurekahedge database

of annualised returns and volatility and so is not that useful in practice since it does not include inter-fund correlations. It is for this reason that another more sophisticated method for visualising clustering analysis and the taxonomy of hedge funds was developed using ultra metric spaces.

6.2.1 Metric Definitions

`Y = pdist(X,metric)` is the MATLAB® command to compute the distance between objects in the data matrix *X*, using the method specified by *metric*. For a data set made up of m objects, there are $m^*(m-1)/2$ pairs in the data set. The result of this computation is commonly known as a distance or dissimilarity matrix (see MATLAB® Statistics User Guide: Section 11-4).

For hedge funds, the following *metrics* are among the most commonly used:

"Euclidian"	Euclidian distance (default)
"City block"	City block distance
"Mahalanobis"	Mahalanobis distance
"Correlation"	One minus the sample correlation between points

where

Euclidian distance: $d_{rs}^2 = (x_r - x_s)(x_r - x_s)'$

City Block distance: $d_{st} = \sum_{j=1}^{n} \left| x_{sj} - y_{tj} \right|$

Mahalanobis distance: $d_{rs}^2 = (x_r - x_s)V^{-1}(x_r - x_s)'$

where V is the sample covariance matrix.
Correlation distance:

$$d_{rs}^2 = \frac{(x_r - \bar{x}_r)(x_s - \bar{x}_s)'}{[(x_r - \bar{x}_r)(x_r - \bar{x}_r)']^{1/2}[(x_s - \bar{x}_s)(x_s - \bar{x}_s)']^{1/2}}$$

where

$$\bar{x}_r = \frac{1}{n}\sum_j x_{rj} \text{ and } \bar{x}_s = \frac{1}{n}\sum_j x_{sj}$$

6.2.2 Creating Dendrograms

Dendrograms are the most commonly used graphical method of taxonomy visualisation and cluster analysis for hedge funds. To make it easier to see the relationship between the distance information generated by `pdist` and the objects in the original data set, the distance vector can be reformatted into a matrix using the `squareform` function. In the squareform matrix, element *I,j* corresponds to the distance between object *I* and object *j* in the data set under analysis. Next, the `linkage` function takes the information generated by `pdist` and links pairs of objects that are close together into binary clusters

(clusters made up of two objects only). Then the `linkage` function links these newly formed binary clusters to each other and to other objects to create bigger clusters until all the objects in the data set under consideration have been linked together into a hierarchical tree or dendrogram.

In a dendrogram, the links between objects are represented as U shaped lines. *The height of the U indicates the distance between the objects*. This height is also known as the *cophenetic* distance.

6.2.3 Interpreting Dendrograms

To better understand exactly what it is a dendrogram represents for each metric, it was decided to run some tests on a data set consisting of 11 time series returns with different Hurst exponents (Mandelbrot 1982) ranging from 0.25 to 0.75 (see hurst_data.xlsx). A random or Gaussian time series scales in time with a volatility which is the square root of the variance. In the same way, a time series shares the same exponent – 0.5 – when it comes to predictability at the special Gaussian halfway mark between the Dimension D = 1 (a straight line) and Dimension D = 0 (a surface). Time series with Hurst exponents greater than 0.5 become progressively elongated (persistent) as the Hurst exponent reaches 1.0. As such they possess trends at many different frequencies and provide positive alternative beta when trend followed as a function of the Hurst exponent. In the same way, time series with Hurst exponents less than 0.5 become more jagged, prone to reversals and are said to be anti-persistent. They too however exhibit “form” or “order”, and this form can also be extracted using trend following but where the opposite buy or sell logic is used (counter trend following).

Regardless of the univariate predictive qualities of each data set as a function of H, their use here will be in establishing an idea of what a dendrogram is showing. Each time series in hurst_data.xls is mutually exclusive and so there is no correlation whatsoever between the data sets. The noise is referred to as fractional Brownian motion. However, there is a clear increase in volatility of each time series as the Hurst exponent decreases as shown in Figure 6.3. The data set is therefore useful in assessing what a dendrogram actually shows since we know there are no correlations but a clear increase in volatility per time series as the Hurst exponent decreases in decrements of 0.05 from 0.75 to 0.25, where 0.5 (GBM) is the middle time series.

Figure 6.4 is the dendrogram generated from the Source code 6.1 with no metric input since the default MATLAB® metric is the Euclidian distance. The distance is shown on the x-axis with the y-axis being the various time series Hurst exponents. As can be seen, the dendrogram is a hierarchical method of checking the clustering tendencies of a complex data set – showing us patterns which would not be obvious otherwise. As can be seen, the Euclidian distance provides what looks like an approximate way of ranking the variance

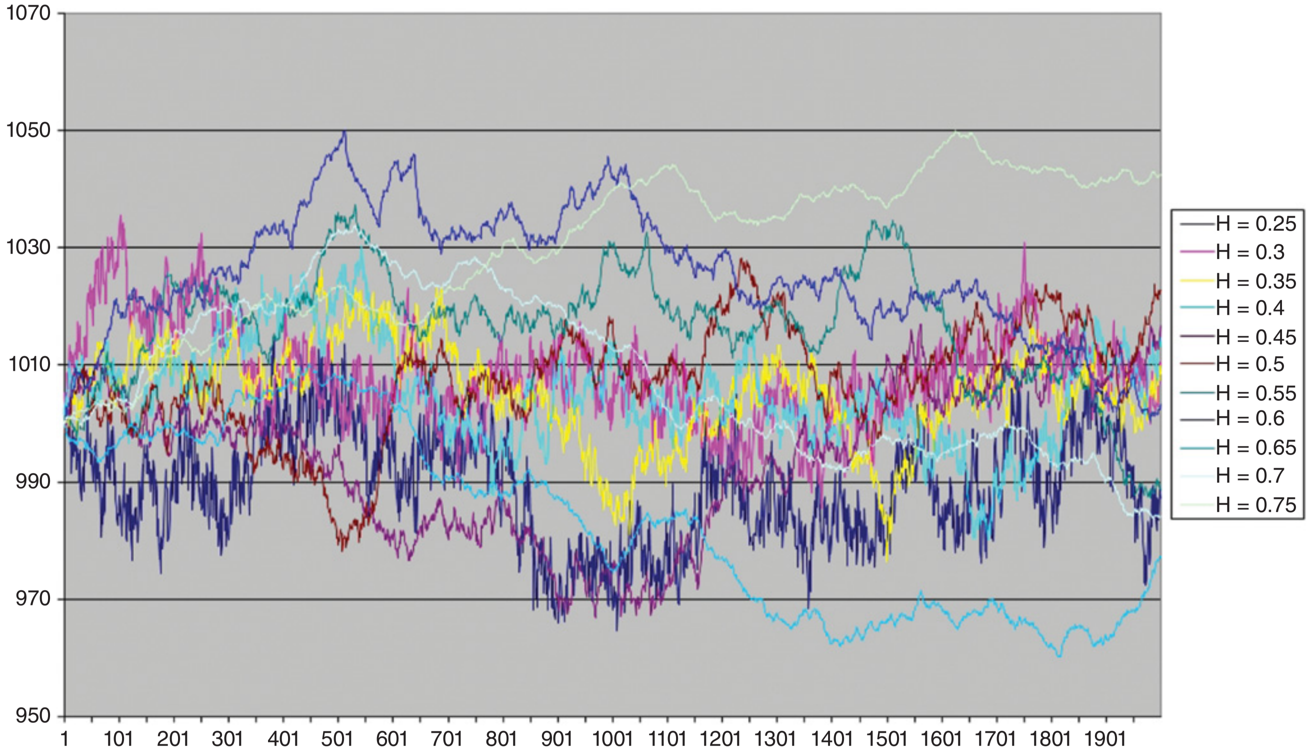


Figure 6.3 Hurst data time series showing how the volatility of each time series increases as the Hurst exponent decreases from 0.75 to 0.25 in decrements of 0.05

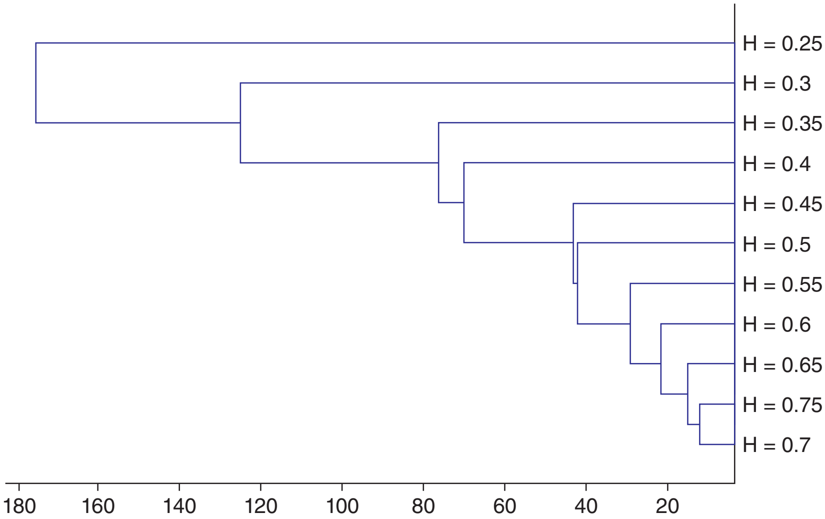


Figure 6.4 Hurst data set Euclidian dendrogram

or volatility since the time series on the y-axis are exactly in order of decreasing volatility going down the axis (increasing Hurst exponent).

Figure 6.5 is the City Block metric dendrogram for the Hurst data. As can be seen it is identical to the Euclidian except that the City Block metric on the x-axis is significantly greater in value. Since we have established that the Euclidian and the City Block metric seem to be more or less the same, we will only consider the Euclidian and not the City Block in various future examples.

```
%File: ddhurst.m
%import XL data
[returns,names] = xlsread('hurst_data.xlsx','H Returns');

names = names(1,1:11);  %get Hurst value labels
returns = returns';  %transpose returns

Y = pdist(returns,'cityblock');  %using City Block distance
squareform(Y);  %reformat distance
Z = linkage(Y);  %link into a hierarchical cluster

%create dendrogram
dendrogram(Z,'colorthreshold',0,'orientation','left','labels',names);
```

Source 6.1 MATLAB® code for creating a dendrogram using a range of metrics

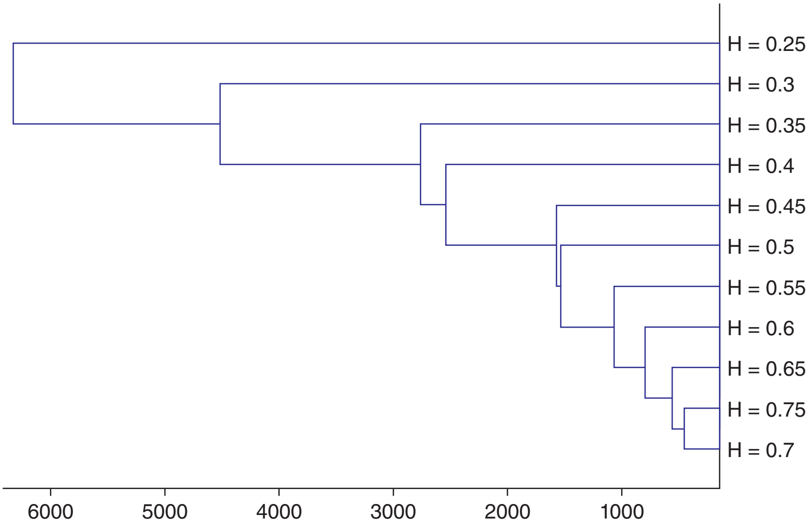


Figure 6.5 City Block dendrogram

Figure 6.6 shows that the Mahalanobis metric produces a result of no real use since the y-axis values are arranged in no particular order of interest. Lhabitant (2004) describes the Mahalanobis distance as a weighted, normalised Euclidian distance where the weighting is determined by the sample variance-covariance matrix with potential uses if the time series are correlated or have

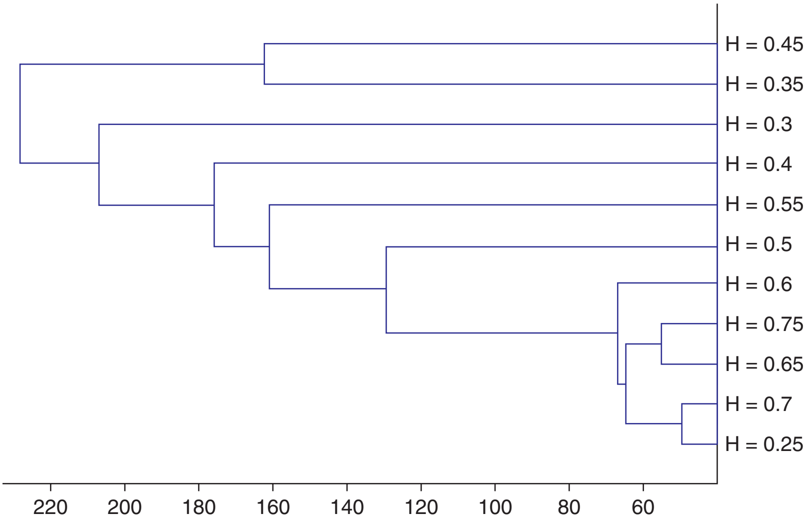


Figure 6.6 Mahalanobis dendrogram

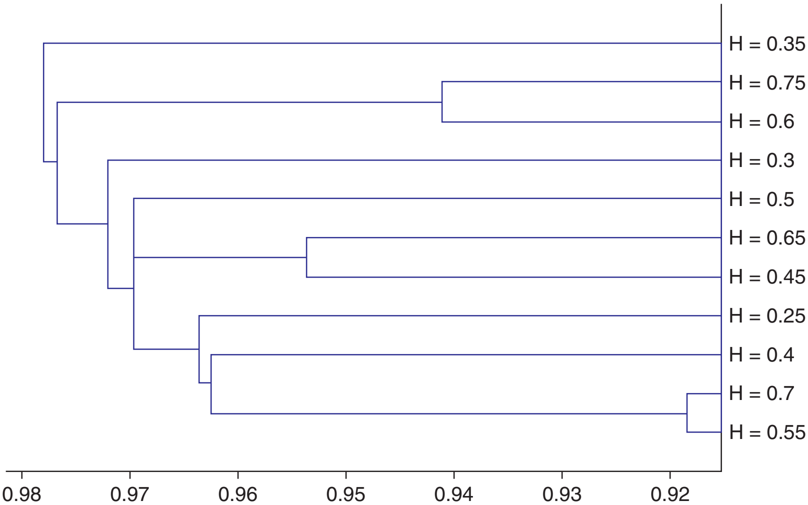


Figure 6.7 Correlation dendrogram

very different variances. Computing the Mahalanobis distance was found to be computationally time-consuming and often generated confusing output for analysis purposes. For this reason, we will not use the Mahalanobis distance in any further examples as it would seem to have a niche application. The Minkowski distance also exists but is rather complicated and obscure – please see the MATLAB® Statistics Toolbox User Guide for further information.

Figure 6.7 shows the dendrogram with the Correlation metric used. As would be expected, there is nothing obvious in the groupings and hierarchies since the data used is not at all correlated. However, from here on in, we will look at examples only using the Euclidian and Correlation metrics. As will be seen, the Correlation metric exposes a wealth of information in real life data sets. In the world of hedge funds leverage is a variable used by the manager and investor alike – and correlation analysis removes leverage (see Equation 5.31). The Correlation metric is therefore particularly useful for hedge fund analysis for this reason, and secondly because the main conclusion of Markowitz was that in a large enough portfolio, the final variance reduction is driven only by the average of the co-variances and not the variances. Correlation is covariance rescaled and as such captures the interesting clusters and hierarchies of data sets which are at the same time de-leveraged to an unleveraged cash state and arranged so that the portfolio manager can see visually the clusters of closely related assets in terms of correlation. These clusters will obviously signal areas of concentration of risk for the portfolio manager since assets with similar correlations will cluster together.

Before we look at some real life data sets, Source 6.2 shows the MATLAB® code for creating the correlation dendrogram of the 10 hypothetical hedge

```
%File: ddtenfundscorrel.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');

names = names(1,2:11);  %get hedge fund names
returns = returns';  %transpose returns
color = input('Input color threshold e.g. 0.5: ');  %input
color threshold

Y = pdist(returns,'correlation');  %using correlation metric
squareform(Y);  %reformat distance
Z = linkage(Y);  %link into a hierarchical cluster

%create dendrogram
dendrogram(Z,'colorthreshold',color,'orientation','left','labels',
names');
```

Source 6.2 MATLAB® code for creating a correlation dendrogram for the 10 hypothetical hedge funds

funds from the Excel file: `hfma_matlab_data.xlsx`. Since the hypothetical data used has been generated from a *stochastic process*[1] through a *Stochastic Differential Equation*[2] (SDE) using *Geometric Brownian*[3] *Motion* (GBM), it will also give random results when seen as a correlation dendrogram which is verified by the random scattering and clustering of the funds on the y-axis with no obvious linkages or correlations (see Figure 6.8). This shows you at least the effects of using white noise in the inputs – randomness in the outputs (or garbage in garbage out).

We know from the initial analysis performed on the Hurst data set that the Euclidian metric as shown in the dendrogram in Figure 6.9 is clustering and ranking the 10 hedge funds approximately according to their volatilities as seen on the y-axis.

Now that we have an idea of what we are looking at when it comes to the Euclidian and Correlation metrics, let's look at some real-life examples. You will see that there are very interesting patterns of clusterings which will emerge from the three data sets considered – real data set 1. 10 continuous futures time series in monthly format, real data set 2. 30 CAC 40 stocks and real data set 3. the style indices of the Eurekahedge database.

As seen in Figure 6.10, the major drawback of the Euclidian approach is that it creates linkages based on volatility levels of the underlying time series.

[1] In a *stochastic process* there is some level indeterminacy in the future evolution of a variable described by a probability distribution. If the initial value is known with certainty, there will still be a variety of possibilities (i.e. randomness) that the process will undertake in the future.

[2] A *Stochastic Differential Equation* (SDE) is an equation for a stochastic process that describes the time evolution of a probability distribution.

[3] Robert Brown (1773–1858) was a Scottish botanist who made important contributions to the field through his pioneering work on the microscope and pollen floating on water.

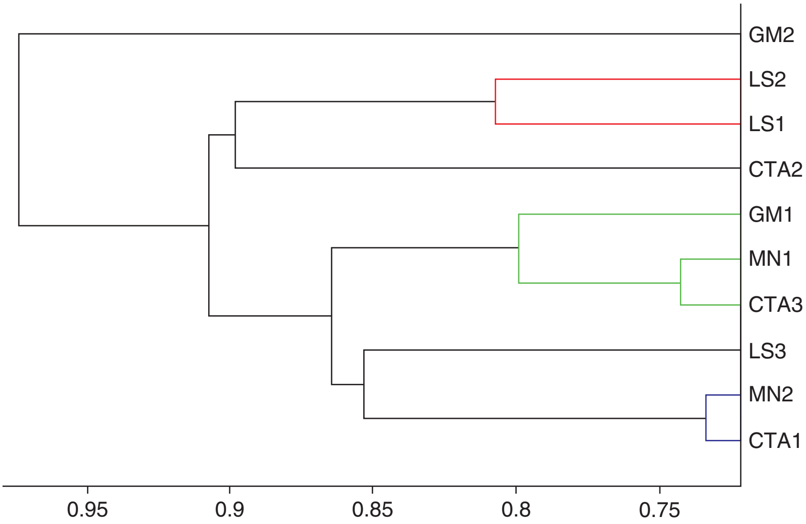


Figure 6.8 Correlation dendrogram for the 10 hypothetical hedge funds

The data set 1 futures example uses tick data in the sense that the historical continuous futures prices have been differenced to find the number of ticks moved each day for each futures market. This is subtly different from finding a return which is not possible using forward adjusted continuous futures data since the values the adjusted futures take in the future bear no resemblance to the actual futures price on that day, so no reference base for percentage

```
%File: ddtenfundseuclidian.m
%import XL data
[returns,names] = xlsread('hfma_matlab_data.xlsx','10 Hedge Funds');

names = names(1,2:11);  %get hedge fund names
returns = returns';  %transpose returns

color = input('Input color threshold e.g. 30: ');  %input color
threshold

Y = pdist(returns);  %using Euclidian distance (default)
squareform(Y);  %reformat distance
Z = linkage(Y); %link into a hierarchical cluster

%create dendrogram
dendrogram(Z,'colorthreshold',color,'orientation','left','labels',
names');
```

Source 6.3 MATLAB® code for creating a Euclidian dendrogram for the 10 hypothetical hedge funds

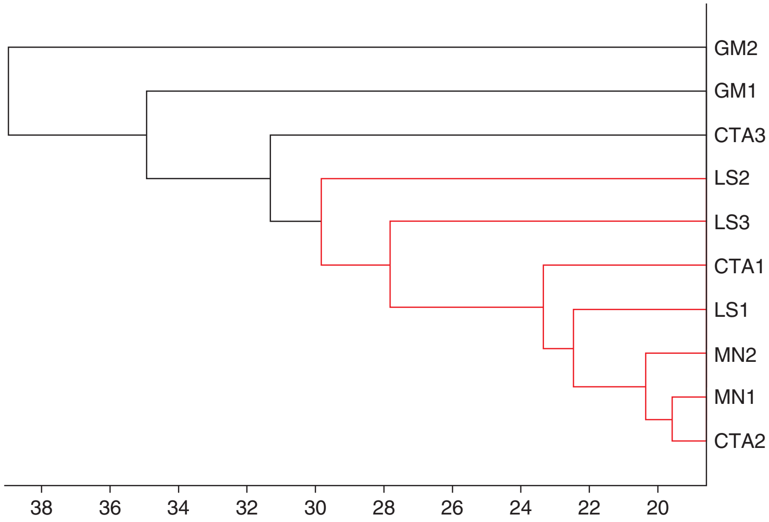


Figure 6.9 Euclidean dendrogram for the 10 hypothetical hedge funds

calculations can be established. Often the CTA uses data like this and so, for this kind of application, the Euclidian metric is clearly of no use as can be seen in Figure 6.10.

Let us now analyse the same data set 1 – monthly futures – using the Correlation metric.

At once, Figure 6.11 stands out in comparison to all previous examples. It is a striking example of the success of the correlation dendrogram as applied to a

```
%File: ddfutureseuclidian.m
%import XL data
[ticks,names] = xlsread('futures_monthly.xlsx','Monthly Ticks');

names = names(1,2:26);  %get futures names
ticks = ticks';  %transpose monthly ticks

color = input('Input color threshold e.g. 140: ');  %input color
threshold

Y = pdist(ticks);  %using Euclidian distance (default)
squareform(Y);  %reformat distance
Z = linkage(Y);  %link into a hierarchical cluster

%create dendrogram
dendrogram(Z,'colorthreshold',color,'orientation','left','labels',
names');
```

Source 6.4 MATLAB® code for creating a Euclidian dendrogram for the monthly futures data

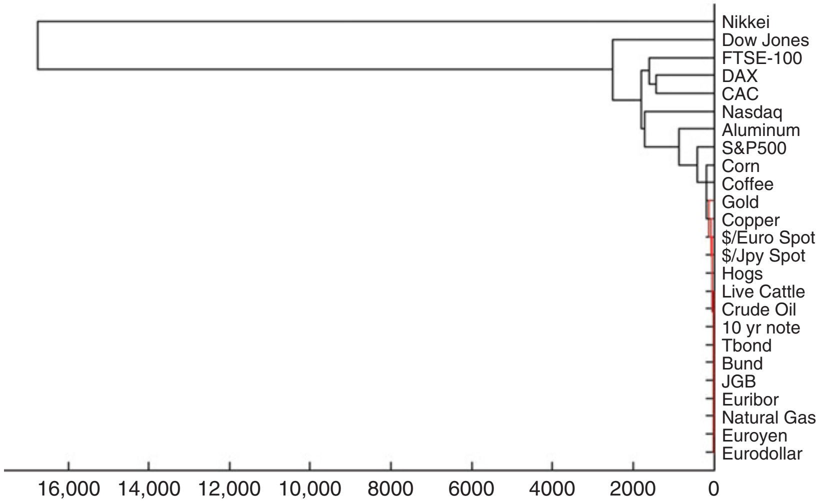


Figure 6.10 Euclidian dendrogram for the monthly futures data

range of liquid futures markets, exactly the kind of underlying instruments and groups used by CTAs and global macro managers and which represent most of the building blocks as reported in Table 6.1. There are clear groupings on the y-axis going up corresponding to the US, European and Japanese bond futures and money markets, next the group of stock indices, then currencies, and finally commodities with the energies, metals and livestock seen correctly grouped. This example shows how a diversified CTA manager sees the world and

```
%File: ddfuturescorrelation.m
%import XL data
[ticks,names] = xlsread('futures_monthly.xlsx','Monthly Ticks');

names = names(1,2:26);  %get futures names
ticks = ticks';  %transpose monthly ticks

color = input('Input color threshold e.g. 0.5: ');  %input color
threshold

Y = pdist(ticks,'correlation');  %using correlation metric
squareform(Y);  %reformat distance
Z = linkage(Y);  %link into a hierarchical cluster

%create dendrogram
dendrogram(Z,'colorthreshold',color,'orientation','left','labels',
names')
```

Source 6.5 MATLAB® code for creating a correlation dendrogram for the monthly futures data

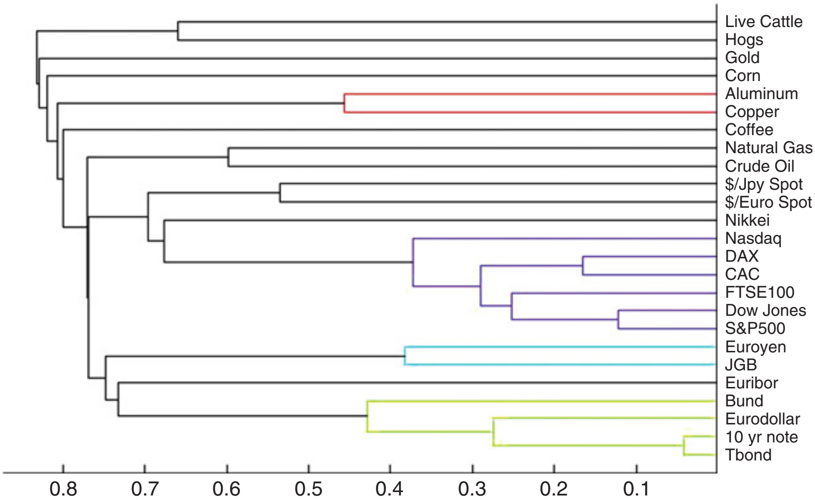


Figure 6.11 Correlation dendrogram for the monthly futures data or "How a Diversified CTA Sees the World"

demonstrates how the dendrogram approach successfully groups the various building blocks into a meaningful clustered hierarchy of correlations.

Figures 6.12 and 6.13 show the dendrograms of 30 stocks out of the CAC 40 (French30) index using the Euclidian and Correlation distances respectively. Again, looking the correlations in Figure 6.13, it is seen that sub-clusters

```
%File: ddfrencheuclidian.m
%import XL data
[returns,names] = xlsread('French30_daily.xlsx','French30 Returns');

names = names(1,2:31);  %get company names
returns = returns';  %transpose returns

color = input('Input color threshold e.g. 100: ');  %input color
threshold

Y = pdist(returns);  %using Euclidian distance (default)
squareform(Y);  %reformat distance
Z = linkage(Y);  %link into a hierarchical cluster

%create dendrogram
dendrogram(Z,'colorthreshold',color,'orientation','left','labels',
names');
```

Source 6.6 MATLAB® code for creating a Euclidian dendrogram for the French30 data

```
%File: ddfrenchcorrel.m
%import XL data
[returns,names] = xlsread('French30_daily.xlsx','French30 Returns');

names = names(1,2:31);  %get company names
returns = returns';  %transpose returns

color = input('Input color threshold e.g. 0.5: ');  %input color
threshold

Y = pdist(returns,'correlation');  %using correlation metric
squareform(Y);  %reformat distance
Z = linkage(Y);  %link into a hierarchical cluster

%create dendrogram
dendrogram(Z,'colorthreshold',color,'orientation','left','labels',
names');
```

Source 6.7 MATLAB® code for creating a correlation dendrogram for the French30 data

are correctly identified going up the y-axis, e.g. car manufacturers, banking, luxury brands, food and beverage manufacturers and so on all the way to the energy sub-cluster at the top. The graphic importance of Figure 6.13 is evident for the Long-Short Equity hedge fund manager who may be trading the 30 stocks as shown in their authorised basket of securities. At a glance the

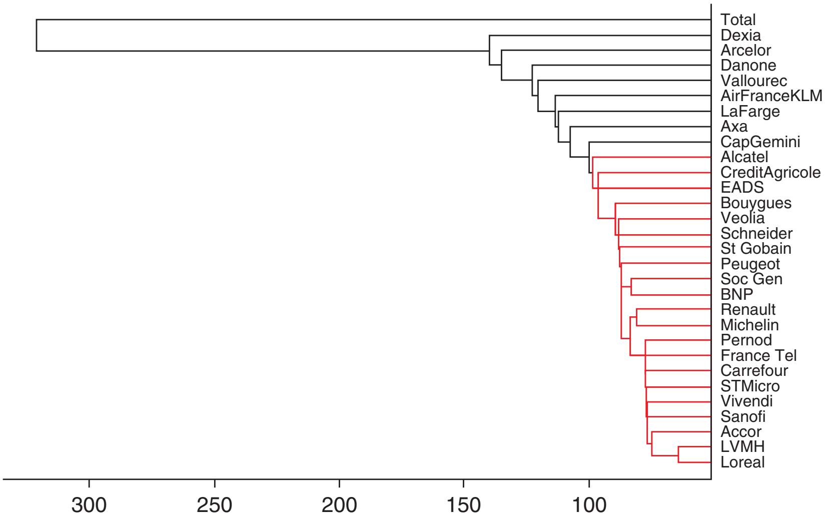


Figure 6.12 French30 Euclidian dendrogram

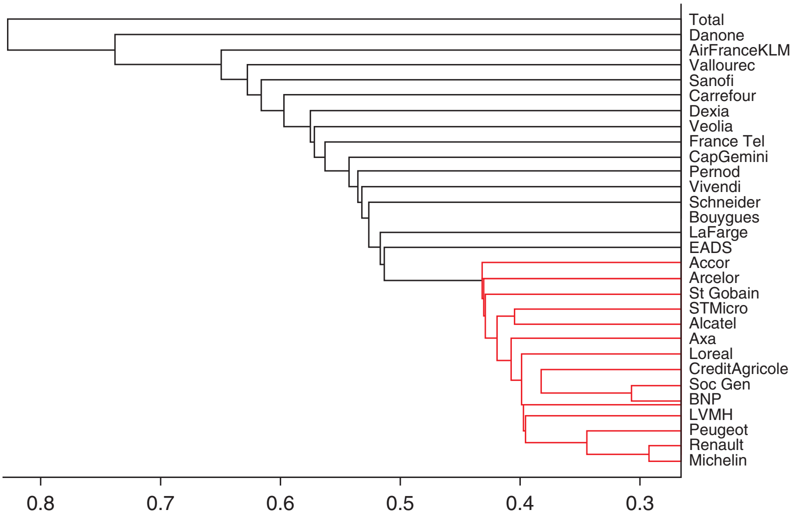


Figure 6.13 French30 correlation dendrogram or "How a French Long-Short Equity Manager sees the World"

manager can see potential groups of concentration of risk or possible pair trade combinations if the manager has, for example, either a divergent or convergent price forecast over time of the stocks in question.

Figure 6.14 is the last example taken from real life i.e. the Eurekahedge data set of style indices from January 1990 to November 2011 using the Correlation metric. It can be seen that again, the Correlation dendrogram accurately

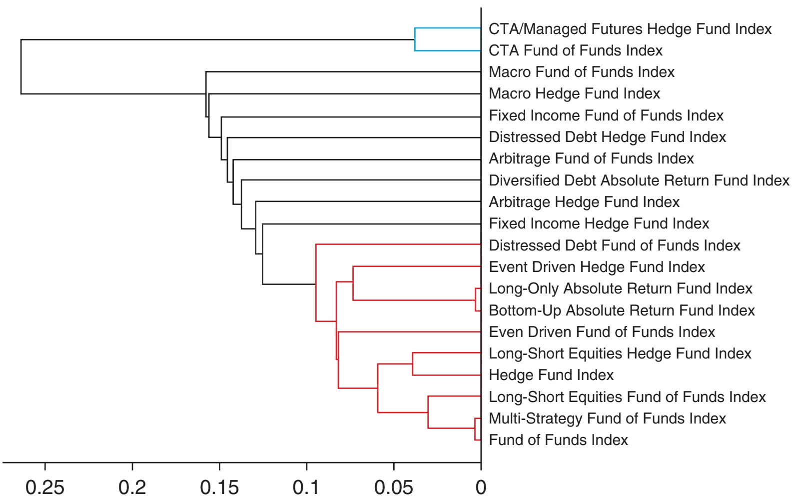


Figure 6.14 Eurekahedge fund correlation dendrogram or "How a Fund of Hedge Funds Manager sees the World"

pinpoints clusters and sub-clusters in the data set. It shows in particular a strong similarity between Multi-Strategy Funds of Funds Index and the Funds of Funds Index, as well as the Long-Only Absolute Return Fund Index and the Bottom-Up Absolute Return Fund Index – as would be expected. There are also interesting clusters for the Long-Short Equities Hedge Fund Index and the Hedge Fund Index, as well as the CTA Managed Futures Hedge Fund Index and the CTA Funds of Funds Index. Most interestingly is the clear departure in terms of cophenetic distance from the CTA cluster to the rest of the set – confirming empirically that CTAs are a style apart from the rest of the hedge fund industry, with significantly different sets of return correlations over time. This is one of the main reasons why CTAs are often included in alternative investment fund of hedge fund portfolios since they have the inherent ability to reduce the total portfolio risk and, by doing so, to increase total risk-adjusted returns of the portfolio.

7 Market Risk Management

We often encounter problems relating to the risk that the value of a hedge fund will decrease due to the impact of various market factors, for example changes in interest or foreign currency rates. Moreover, with the heightened publicity of recent financial events, hedge fund managers have come under increased pressure from investors and regulators to efficiently manage, monitor, measure and report such market risk inherent in their investment strategies. Indeed, experience has clearly shown that the measurement and management of extreme market conditions is of paramount importance for hedge funds.

Chapter 7 provides an introduction to market risk management for hedge funds and presents the fundamentals of quantitative risk measures and models used in the industry today. The chapter also covers some of the more advanced risk measures available that can more effectively manage risk in a hedge fund in light of the limitations encountered with traditional market risk measures.

7.1 VALUE-AT-RISK

The Value-at-Risk (VaR)[1] for a portfolio of assets is the worst estimated loss over a given time horizon (e.g. monthly) at a specified level of confidence (e.g. 95%). That is, the riskiness of the hedge fund portfolio at a specific level of probability in the *left-tail* of the P&L distribution. VaR is often based on the assumption that asset returns follow a *normal distribution* and that the performance of the hedge fund portfolio is affected by a set of linear market factors. As discussed in Chapter 3, under such assumptions it is possible to describe the distribution of asset returns by just two statistical parameters, i.e. μ and σ. That is, assuming that a distribution of monthly hedge fund P&Ls are characterised by a normal distribution, then, at a confidence, $c\%$, there is an expected loss for the hedge fund of no more than z_α standard deviations (i.e. $z_\alpha \sigma$) below the mean over the next month.

[1] VaR is an example of a *downside risk* measure, i.e. the likelihood that an investment will decline in value, or the amount of loss that could result from such a potential decline.

Box 7.1 Quantiles and Percentiles

Quantiles are points taken at regular intervals from a probability distribution. More formally, the quantile function for any probability distribution is the inverse of the Cumulative Distribution Function (CDF). Some q-quantiles have special names, for example:

- 4-quantiles are known as quartiles
- 100-quantiles are known as percentiles

For an ordered set of data from smallest to largest (i.e. an ascending order) the required percentile is a value that represents the number below which a certain percentage of the data fall, e.g. the 5th percentile is the value below which 95% of all the observations fall. The built-in MATLAB® functions `prctile()` or `quantile()` can be used to achieve the relevant values at the desired percentage or cumulative probability, respectively.

In MATLAB® percentiles are specified using percentages ranging from 0 to 100. *Linear interpolation* is used to determine percentiles for percentage values between 100*(0.5/`N`) and 100*((`N`-0.5)/`N`). Note that this can lead to misleading results if `N` is not exactly 100.

For example, at a 95% confidence level (5th percentile or significance), 95% of the time the loss is expected to be no worse than 1.645σ. That is, the critical value is –1.645 indicating that there is a 5% probability that a particular value will be at least 1.645 standard deviations below the mean (i.e. -1.645σ) (see Figures 7.1 and 7.2). The critical value is found using the built-in MATLAB® function `norminv()` which returns the inverse of the CDF function. Source 7.2 shows the calculation of the two most common confidence levels at 95% and 99% along with their associated critical values z_α for the standard normal distribution.

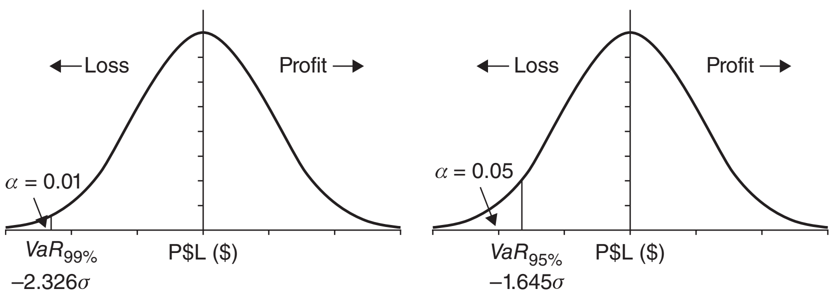


Figure 7.1 VaR at the 95% and 99% confidence levels

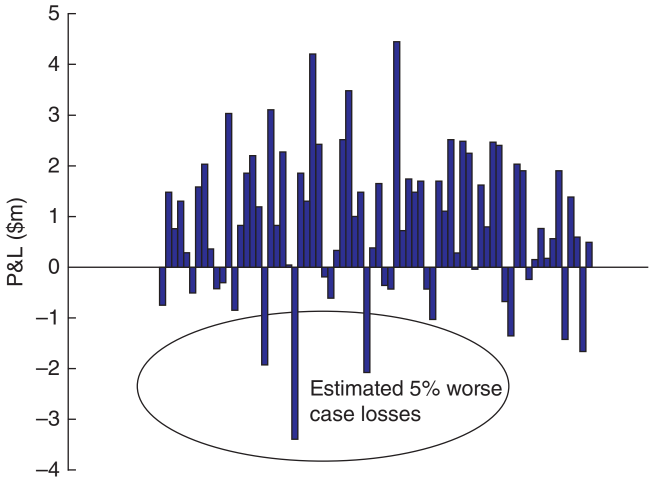


Figure 7.2 Estimated 5% worse case losses for a hypothetical CTA index (2008–2013)

For a standard normal distribution (i.e. $\mu = 0$ and $\sigma = 1$), the VaR for a hedge fund at a 100(1- α)% confidence level, $VaR_{1\text{-}\alpha}$ is given by:

$$VaR_{1-\alpha} = Z_{\alpha} \tag{7.1}$$

```
%File: pnlchart.m
%import XL data
[~,dates,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

N = size(returns,1); %# data points
AuM = 100; %$100,000,000 hedge fund value
PnL = (returns./100).*AuM; %calculate monthly P&L ($)

figure; %create figure
sDate = datenum(dates(1)); %set start date for x-axis
eDate = datenum(dates(end)); %set end date for x-axis
xData = linspace(sDate,eDate,N);

bar(xData, PnL); %plot P&L chart
box off;
set(gca, 'XTick', []); %format x-axis
set(gca,'XColor','w'); %format y-axis
ylabel('P&L ($m)');
```

Source 7.1 MATLAB® code for the P&L chart in Figure 7.2

```
%File: critical.m
%VaR
z5 = norminv(0.05); %5% significance (alpha) i.e. 95% confidence level
z1 = norminv(0.01); %1% significance (alpha) i.e. 99% confidence level

OUTPUT:

Critical values (VaR)
z5 = -1.645
z1 = -2.326
```

Source 7.2 Critical values at 95% and 99% confidence levels

where Z_α is the critical value from the standard normal distribution at the required significance level, α.

So, VaR at the $100(1 - \alpha)\%$ confidence level is defined as the lower $100\alpha^{th}$ percentile of the P&L distribution (see Box 7.1). However, since μ and σ are parameters[2] of the hedge fund P&L distribution, $VaR_{1-\alpha}$ can be written more formally as:[3]

$$VaR_{1-\alpha} = \mu + Z_\alpha \sigma \tag{7.2}$$

Note that $VaR_{1-\alpha}$ scales with the volatility (see Box 7.2). VaR is usually reported on a monthly basis in negative dollar terms, which further emphasises that it is a measure of losses, or as an *absolute* positive dollar amount. In Equation 7.2, since Z_α is negative, $VaR_{1-\alpha}$ will always be a negative value indicating a dollar loss. Some fund managers report VaR on a monthly basis in percentage terms so as to be consistent with the reporting of other risk measures. Note that VaR does not give any information about the amount of loss (i.e. the magnitude) expected in excess of VaR but only indicates that $100\,\alpha\%$ of the time the loss to the hedge fund is estimated to be at least as bad over a certain period.

Box 7.2 Square Root Rule – VaR

If a series of hedge fund P&Ls are quoted in monthly or quarterly figures then they can be transformed into an equivalent annualised series using a similar *square root rule* to that applied to standard deviation. To get the annualised figure, the original VaR is multiplied by the square root of the frequency representing the original time period, e.g. 12 for monthly and 4 for quarterly.[4] More formally:

$$\text{annual } VaR_{1-a} = \text{monthly } VaR_{1-a} \times \sqrt{12}$$
$$\text{annual } VaR_{1-a} = \text{quarterly } VaR_{1-a} \times \sqrt{4}$$

[2] Indeed, this is an example of the parametric method for calculating VaR.

[3] In Equation 7.2, μ is used to centre the normal distribution, before subtracting the relevant number of σ's to get the VaR.

[4] When considering daily P&Ls, practitioners generally assume there are 252 trading days in a year.

7.2 TRADITIONAL VaR METHODS

There are generally three industry accepted methods for estimating VaR, namely:

1. Historical simulation,
2. Parametric method, and
3. Monte-Carlo simulation.

Each has a different approach in terms of how they describe the distribution of losses. Monte-Carlo simulates data, historical simulation uses actual data, and the parametric approach utilises the data but only in order to generate the necessary parameters to characterise the distribution. All of these traditional measures of VaR have their strengths and weaknesses.

7.2.1 Historical Simulation

Of all the traditional VaR methods, the historical simulation (or non-parametric) method is probably the simplest to implement since only a set of historical hedge fund returns over a given time period is required. In general, the historical simulation method requires a relatively long history of returns in order to get a meaningful value for VaR. That is, determining VaR for a hedge fund with only a few data points will not provide a good estimate of VaR. It is important to be careful not to draw any conclusions from the data if it is not large enough to be representative of the returns distribution in the future. In this sense, the assumption that historic monthly hedge fund returns are an accurate representation of the future is a major disadvantage of this method since there is no certainty that the past will replicate the future.

Estimating VaR by historical simulation involves calculating a series of simulated P&L values based on a set of historical hedge fund returns. A set of simulated P&Ls are generated for each hedge fund return and ordered ascending (i.e. smallest to largest) using the built-in MATLAB® function `sort()`. For a 95% confidence level (e.g. $\alpha = 0.05$), the $100\alpha = 5th$ percentile in the sorted P&L is used to estimate VaR. That is, if we want to know the worse expected loss at 95% confidence then we look at the 5th percentile since, on an historical basis, the simulated P&L distribution tells us that 95% of the time we will not lose any more than this amount. Source 7.4 shows the MATLAB® code for an implementation of the historical simulation method to estimate the VaR at a 95% confidence level for a hypothetical CTA Index of monthly returns with a value of \$100m. The historical monthly VaR is estimated at \$1.645m for a 95% confidence level, i.e. over the next month there is a 5% probability that the CTA Index will lose \$1.645m in value.[5] In order to express VaR in percentage terms,

[5] With historical simulation, since *actual* monthly returns are used, the distribution is already assumed to be centred, so there is no need to subtract the relevant number of σ's from μ in order to get the VaR figure.

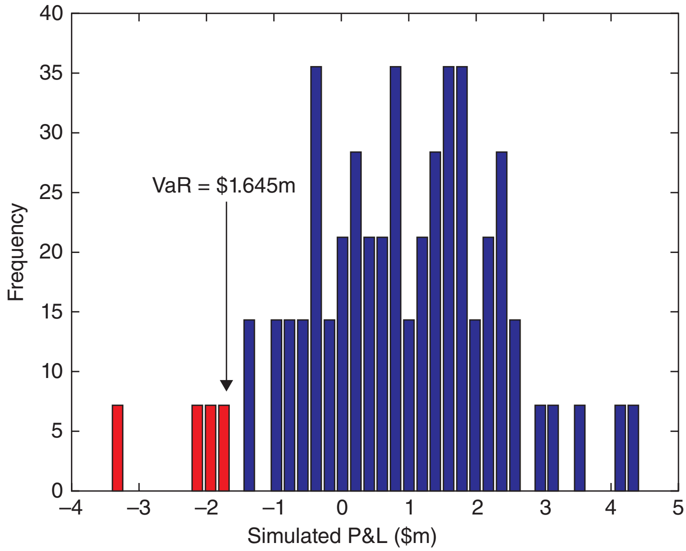


Figure 7.3 Histogram showing monthly VaR using historic simulation

simply divide by the current value of the hedge fund (i.e. $100m). Figure 7.3 shows the distribution of monthly P&L values and the VaR at 95% confidence level showing a loss in the left-tail of the distribution at the 5th percentile. Note that there are losses in excess of the VaR value further to the left of the tail which indicates that VaR does not tell us anything about the actual magnitude of the extreme losses in the tail.

```
%File: historic.m
%import XL data
[~,~,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

P = 100; %$100m hedge fund value
PnL = (returns./100).*P; %calculate simulated monthly P&L ($m)
ptile = 5; % 95% confidence level

srtPnL = sort(PnL); %sort P&L smallest to largest
mhVaR = prctile(srtPnL,ptile); %monthly VaR ($m)

OUTPUT:

monthly VaR ($m) = -1.645
```

Source 7.3 MATLAB® code for the calculation of the monthly VaR at 95% confidence level using historical simulation

```
%File: historic.m
%<....>

%Plot histogram
N = size(returns,1); %# data points
[count,bins] = hist(returns,40);
scale = (bins(2)-bins(1))*N; %scale bins
fcount = count/scale; %frequency (count)
cutoff = count.*(bins<mhVaR); %P&L<VaR (i.e. 5% tail cut-off)
fcutoff = cutoff/scale; %frequency (cut-off)

figure;
bar(bins,fcount*100,'b'); %plot full data set
hold on;
bar(bins,fcutoff*100,'r'); %plot cut-off data
hold off;
xlabel('Simulated P&L ($m)'); %format axes
ylabel('Frequency');
annotation('textarrow',[.33,.33],[.6,.28],'String','VaR = $1.645m');
%add arrow & text
```

Source 7.4 MATLAB® code for the histogram plot in Figure 7.3

7.2.2 Parametric Method

The parametric method does not require all of the data that the historic simulation requires but may use the data to determine parameters of the hedge fund return distribution. If we generally assume the hedge fund returns are consistent with a normal distribution then we only need two parameters to characterise the distribution, namely the mean, μ and standard deviation, σ. That is, once we have determined those two parameters we no longer require all of the historical data to calculate VaR. Recall from Equation 7.2 that the VaR can be written formally as:

$$VaR_{1-\alpha} = \mu + Z_{\alpha}\sigma \tag{7.3}$$

Source 7.5 shows a user-defined MATLAB® function for calculating the VaR for any reporting frequency and significance level, α. Source 7.6 shows an implementation of this function to determine the monthly VaR at a 95% confidence for a hypothetical CTA Index (2008–2013) with a current market value of $100m.

Figure 7.4 shows that the monthly VaR for a CTA Index valued at $100m for a 95% confidence level is $1.573m, i.e. there is only a 5% chance that the value of the CTA Index will fall by more than $1.573m over the next month.

```
%File: fVaR.m
%VALUE-AT-RISK (VaR)
function VaR = fVaR(x,f,alpha)
%------------------------------------------------------------------
%x     - returns
%f     - reporting frequency
%alpha - significance level e.g. 0.05
%VaR   - Value-at-Risk
%------------------------------------------------------------------

   m1 = fMean(x,f); %mean (%)
   m2 = fStd(x,0,f); %std. dev. (%)
   z = norminv(alpha); %critical value
   VaR = m1+m2.*z; %VaR
end
```

Source 7.5 User-defined MATLAB® function for calculating VaR

7.2.3 Monte-Carlo Simulation

The Monte-Carlo (MC) method assumes that a series of hedge fund returns can be characterised by a stochastic (or probabilistic) model. MC methods are a widely used class of computational algorithms for simulating the behaviour of various physical and mathematical systems having been popularised by

```
%File: parametric.m
%import XL data
[~,~,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

P = 100; %$100m hedge fund value
PnL = (returns./100).*P; %calculate simulated monthly P&L ($m)

%if monthly returns -> reporting frequency f = 1 for monthly values
f = 1;

alpha = 0.05; %95% confidence level
mpVaR = fVaR(PnL,f,alpha); %monthly VaR ($m)

OUTPUT:

monthly VaR ($m) = -1.573
```

Source 7.6 MATLAB® code for the calculation of the monthly VaR at 95% confidence level using the parametric method

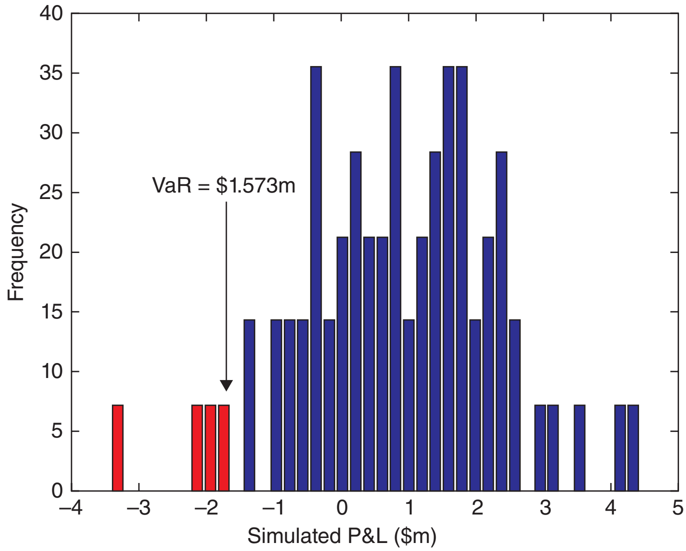


Figure 7.4 Histogram showing monthly VaR using the parametric method

John Von Neumann[6] and Nicholas Metropolis[7] to name but a few. MC are distinguished from other simulation-based methods by being of a stochastic nature, that is the MC model includes a non-deterministic component that introduces a degree of uncertainty or randomness into the process through the use of random number generators.[8] The fundamental idea behind the technique involves simulating thousands of trials (or paths) that a series of hedge fund returns are likely to follow over a certain time period in the future based on a specific stochastic (or probabilistic) model. Each trial leads to a terminal value for the hedge fund P&L at the end of each simulation period. After thousands of such runs, a simulated P&L distribution is obtained from which a VaR can be estimated at a preferred confidence level in much the same way as the historical simulation, i.e. ordering the P&L and locating the relevant percentile P&L value, i.e. loss. A drawback (although not a restrictive one) to the MC method is that the simulated P&L distribution relies on specific model parameters that a series of hedge fund returns are expected to be governed by in the future. Such a model is primarily driven by the mean, μ and standard deviation, σ

[6] John von Neumann (1903–1956) was a Hungarian-American mathematician who made major contributions to a vast range of fields, including set theory, functional analysis, quantum mechanics, economics, game theory, computer science and numerical analysis.

[7] Nicholas Metropolis (1915–1999) was a Greek American physicist.

[8] MC simulations rely heavily on the sampling method and stability of the Random Number Generator (RNG). For this reason, many financial houses spend a great deal of time, money and effort developing better and more robust RNGs.

Table 7.1 Comparison of traditional VaR methods

Method	Advantages	Disadvantages
Historical Simulation	No assumptions about the return distributions.	Assumes data used in the simulation are representative of the future.
Parametric Method	Mathematically simple to understand and implement.	Strong assumptions about the hedge fund P&L distribution in terms of μ and σ. Less accurate for nonlinear instruments used within the hedge fund portfolio and strategy.
Monte-Carlo Method	Flexibility in terms of choosing the stochastic process and allows for the inclusion of subjective judgements into the model.	Most demanding in terms of computational resources. Can become mathematically complex and challenging.

of the hedge fund return distribution determined from historical data as well as the inclusion of a degree of subjective knowledge (i.e. market experience) into the model where necessary. The stochastic model is often the fundamental building block to many MC simulations being used extensively throughout the financial markets. For this reason the MC method is a very powerful and much used technique for estimating VaR within the hedge fund community. Not only is the method robust and probabilistically strong, it is also an excellent way of building non-linearity into the return distribution and facilitating a better understanding of the characteristics of the use of derivatives within the portfolio with greater confidence. However, the MC method can become mathematically challenging and computationally intensive when attempting to model a particular hedge fund portfolio or strategy.

Table 7.1 gives a brief summary of the advantages and disadvantages of the three traditional methods of determining an estimate for VaR.

In addition to the estimation of VaR through either of the traditional methods, hedge fund managers will also carry out a variety of *stress tests* on the hedge fund portfolio. That is, the parameters and risk factors that affect hedge fund performance will be greatly magnified, for example raising the volatility over a particular period (e.g. by 100% or 200%) of the original value so as to cause a serious risk to the hedge fund of losing a catastrophic amount of money. This helps the fund manager understand where problems may be concentrated and allows them to be prepared for such events should they arise (however unlikely). Similarly, fund managers may run *scenario analyses* using a set of historical data and related parameters that cover a specific turbulent period in the financial markets, such as the recent financial crisis of 2008 or the ongoing European economic crisis. This will also help the manager identify potential areas of large losses and allow them to develop strategic measures to alleviate such problems in the event of a similar financial disaster.

7.3 MODIFIED VaR

Despite the use of stress tests and scenario analysis, the most erroneous (and potentially damaging) assumption when using traditional VaR methods is that of hedge fund returns following a normal distribution. Clearly, this is invalid since it is well known that hedge fund returns generally have fatter tails and an asymmetric return distribution, i.e. the presence of negative skewness and positive excess kurtosis. In order to address this issue many extensions to the traditional VaR methods have been put forward as better estimators of hedge fund market risk. Such methods either explicitly incorporate skewness and kurtosis into the model or focus primarily on the left-tails of the returns distribution where most of the extreme negative returns (i.e. large losses) occur.[9] One such extension already discussed in Chapter 5 for the calculation of the modified Sharpe ratio (MSR) is the modified VaR (MVaR). MVaR explicitly takes into account the third and fourth moments of the return distribution, namely skewness and kurtosis. That is, there is a branch of mathematics that is involved with power series expansions of quantile functions (see Box 7.1) such as those related to VaR. Indeed, the higher moments of the distribution are incorporated into the VaR measure using the celebrated Cornish-Fisher expansion (1937), such that the following power series can be obtained for the first few terms:

$$z_{cf} \approx z_\alpha + \frac{1}{6}\left(z_\alpha^2 - 1\right)s + \frac{1}{24}\left(z_\alpha^3 - 3z_\alpha\right)k - \frac{1}{36}\left(2z_\alpha^3 - 5z_\alpha\right)s^2 \tag{7.4}$$

where z_{cf} is the Cornish-Fisher critical value from the normal distribution at the respective significance level, α.

Note that when the return distribution is normally distributed, s and k will both be zero and therefore:

$$z_{cf} \approx z_\alpha \tag{7.5}$$

The MVaR at a confidence level, 100 (1- α)%, $MVaR_{1-\alpha}$ is given by:

$$MVaR_{1-\alpha} = \mu + Z_{cf}\sigma \tag{7.6}$$

where μ is the mean of the returns, s is the sample skewness, k is the sample excess kurtosis and z_α is the critical value from the normal distribution at the required significance level, α.

[9] Many other distributions exist that offer better estimates of VaR for hedge funds, e.g. Johnson distributions and simulated skewed Student's t-distribution.

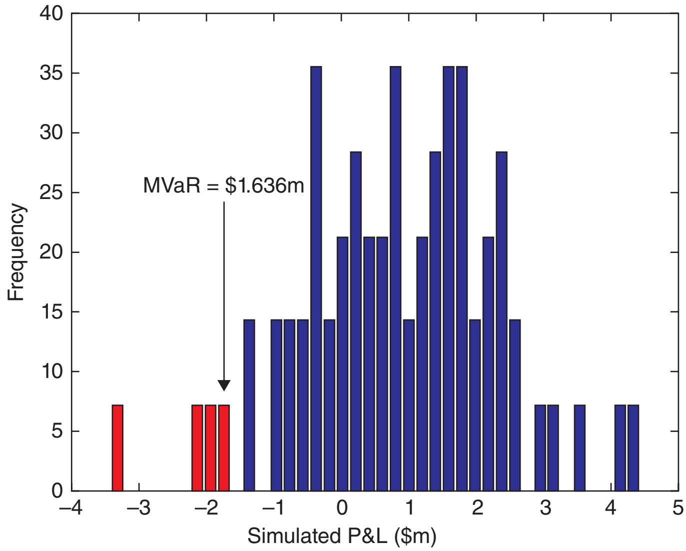


Figure 7.5 Histogram plot showing monthly MVaR

It turns out that using Equation 7.6 leads to a more accurate measure of the VaR, however, there are limitations to this estimation. The higher the confidence level (e.g. 99%) the deeper we are in the left-tail of the distribution and this can lead to inaccurate results since the Cornish-Fisher expansion breaks down at these levels. MVaR is also unreliable when the returns distribution is highly skewed and fat-tailed. Figure 7.5 shows that the monthly $MVaR_{95}$ is $1.636m and clearly higher when compared to the monthly parametric VaR_{95} estimate of $1.573m for the hypothetical CTA Index over the period 2008 to 2013.

7.4 EXPECTED SHORTFALL

Apart from the assumption that the P&L distribution is normal, VaR also fails to satisfy one of the concepts of a coherent risk measure. Artzner *et al* (1999) have stated that a desirable measure of risk should satisfy four basic properties or axioms of risk, namely:

1. Must be monotonic, i.e. if asset $X \geq 0$, $\text{VaR}(X) \leq 0$ i.e. positive returns should not increase risk.
2. Must be sub-additive, i.e. for assets X_1 and $X2$, $\text{VaR}(X_1 + X_2) \leq \text{VaR}(X_1) + \text{VaR}(X_2)$, i.e. the risk of a portfolio of two assets should not be larger than the risk of the sum of the individual assets. If this were the case then adding assets to a portfolio to reduce risk through diversification would not be possible.

```
%File: fMVaR.m
%MODIFIED VALUE-AT-RISK (MVaR)
function MVaR = fMVaR(x,f,alpha)
%---------------------------------------------------------------------
%x      - returns
%f      - reporting frequency
%alpha - significance level e.g. 0.05 (95%)
%MVaR  - Modified Value-at-Risk
%---------------------------------------------------------------------

    m1 = fMean(x,f); %mean (%)
    m2 = fStd(x,0,f); %std. dev. (%)
    m3 = fSkew(x,0,f); %skew (s)
    m4 = fKurt(x,0,f); %excess kurtosis (k)

    z = norminv(alpha); %critical value
    zcf = z+(((z.^2-1).*m3)./6)+(((z.^3-3.*z)*m4)./24)-  ...
          ((((2*z.^3)-5.*z)*m3.^2)./36); %Cornish-Fisher variate
    MVaR = m1+m2.*zcf; %MVaR
end
```

Source 7.7 User-defined MATLAB® function for calculating MVaR

```
%File: MVaR.m
%import XL data
[~,~,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

P = 100; %$100m hedge fund value
PnL = (returns./100).*P; %calculate monthly P&L ($m)

%if monthly returns -> reporting frequency f = 1 for monthly values
f = 1;

alpha = 0.05; %95% confidence level
mMVaR = fMVaR(PnL,f,alpha); %monthly MVaR ($m)
mMVaRp = -(mMVaR/P)*100; %monthly MVaR (%)

OUTPUT:

monthly MVaR ($m) = -1.636
```

Source 7.8 MATLAB® code for the calculation of the monthly MVaR at a 95% confidence level

3. Must possess positive homogeneity so for any positive real number, a, VaR(aX) = aVaR(X), i.e. increasing the size of the portfolio by a times should increase the risk by a multiple of a assuming all the assets within the portfolio remain the same in terms of weighting.
4. Must be translational invariant such that for any real number, a, VaR(X + a) ≤ VaR(X) – a, i.e. adding amount of cash (or risk-free asset) to the portfolio should result in a reduction of the risk by an amount a.

Unfortunately VaR only satisfies three of the axioms of risk, i.e. it fails to satisfy the sub-additive rule. For this reason, an alternative (and often complementary) measure of VaR was developed known as Expected Shortfall (ES)[10] which is discussed in detail in Rockafellar and Uryasev (2000). ES does satisfy all the axioms of risk and is considered a coherent risk measure. ES is the *conditional expectation* of loss given that the loss is beyond VaR, i.e. the loss conditional on that exceeding VaR, that is:

$$ES_{1-\alpha} = E\left[-X \mid -X \geq VaR_{1-\alpha}\right] \tag{7.7}$$

where $ES_{1-\alpha}$ is the estimated ES at a 100(1- α)% confidence level.

$E[X \mid B]$ is the conditional expectation of the random variable X given event B, i.e. the average loss when VaR is exceeded. Yamai and Yoshiba (2002) show that Equation 7.7 can be derived as:

$$\begin{aligned} ES_{1-\alpha} &= E\left[-X \mid -X \geq VaR_{1-\alpha}\right] = \frac{E\left[-X \cdot I_{X \leq -VaR_{1-\alpha}}\right]}{\alpha} \\ &= -\frac{1}{\alpha\sigma_X\sqrt{2\pi}}\int_{-\infty}^{-VaR_{1-\alpha}} t \cdot e^{-\frac{t^2}{2\sigma_X^2}}\,dt = -\frac{1}{\alpha\sigma_X\sqrt{2\pi}}\left[-\sigma_X^2 e^{-\frac{t^2}{2\sigma_X^2}}\right]_{-\infty}^{-VaR_{1-\alpha}} \\ &= \frac{\sigma_X}{\alpha\sqrt{2\pi}} e^{-\frac{VaR_{1-\alpha}^2}{2\sigma_X^2}} = \frac{\sigma_X}{\alpha\sqrt{2\pi}} e^{-\frac{q_\alpha^2\sigma_X}{2\sigma_X^2}} = \frac{e^{-\frac{q_\alpha^2}{2}}}{\alpha\sqrt{2\pi}}\sigma_X \end{aligned} \tag{7.8}$$

where I_A is the indicator function whose value is 1 when A is true and 0 when A is false, and q_α is the 100α percentile of the normal distribution.

When the P&L distribution is normal, VaR does not have the problems pointed out by Artzner *et al* (1999) and as such VaR does not have the issue of left-tail risk. That is, ES and VaR are simply scalar multiples of each other, because they are scalar multiples of the standard deviation, σ. For example, VaR at the 95% confidence level is –1.645σ, while ES at the same confidence level is –2.063σ as shown in Figure 7.6.

[10] Also known as Conditional VaR (CVaR), mean excess loss beyond VaR or tail VaR.

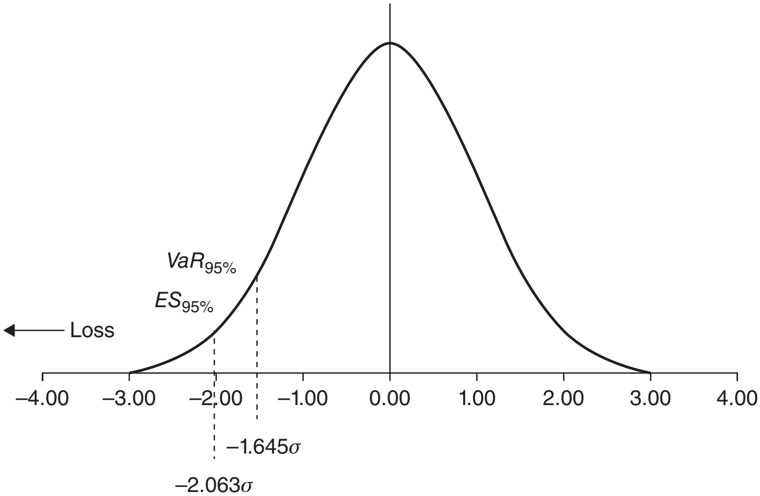


Figure 7.6 Comparison of ES and VaR at the 95% confidence level

The formulation in Equation 7.7 can also be interpreted in terms of the built-in MATLAB® functions `normpdf()` and `norminv()`. Source 7.9 shows the calculation of the ES critical values using the derived formula. Note that both zES5 and zES1 have been negated in both calculations to be consistent with the critical values for VaR and to indicate a loss in the left-tail of the P&L distribution.

Using the standard formula for $VaR_{1-\alpha}$ from Equation 7.2, we have similarly:

$$ES_{1-\alpha} = \mu + Z_{ES\alpha}\sigma \tag{7.9}$$

where $Z_{ES\alpha}$is the critical value from a normal distribution at the required significance level, α.

This gives the estimate of the ES in the left-tail of the P&L distribution but of course it is only the average of the left-tail; the loss could be further along the distribution. Since ES is by definition more than VaR, ES is a more conservative estimate of VaR and thus why it is often used as a complement to traditional VaR measures. Source 7.10 shows a user-defined MATLAB® function for calculating the ES at any reporting frequency and significance level, α.

Source 7.11 shows the calculation of the monthly ES at the 95% confidence level for the hypothetical CTA Index. Clearly, ES is a much larger value than VaR which is to be expected and a much more conservative value to that of traditional VaR measures, offering a greater insight into the actual loss that could be faced over the next month in the value of the CTA Index. Figure 7.7

```
%File: critical.m
%ES
%Using derived formula
%Using derived formula
zES5 = -(exp(-norminv(0.95)^2/2))/(0.05*sqrt(2*pi())); %5%
significance (alpha) i.e. 95% confidence level
zES1 = -(exp(-norminv(0.99)^2/2))/(0.01*sqrt(2*pi())); %1%
significance (alpha) i.e. 99% confidence level

OUTPUT:

Critical values (ES)
zES5 = -2.063
zES5 = -2.665

%Using MatLab functions
zES5 = -normpdf(norminv(0.05))./(0.05); %5% significance (alpha)
i.e. 95% confidence level
zES1 = -normpdf(norminv(0.01))./(0.01); %1% significance (alpha)
i.e. 99% confidence level

OUTPUT:

Critical values (ES)
zES5 = -2.063
zES5 = -2.665
```

Source 7.9 ES critical values at 95% and 99% confidence levels

```
%File: fES.m
%EXPECTED SHORTFALL (ES)
function ES = fES(x,f,alpha)
%----------------------------------------------------------------------
%x     - returns
%f     - reporting frequency
%alpha - significance level e.g. 0.05 (95%)
%ES    - Expected Shortfall
%----------------------------------------------------------------------

        m1 = fMean(x,f); %mean (%)
        m2 = fStd(x,0,f); %std. dev. (%)
        zES = -normpdf(norminv(alpha))./(alpha); %critical value
        ES = m1+m2.*zES; %ES
end
```

Source 7.10 User-defined MATLAB® function for calculating ES

```
%File: ES.m
%import XL data
[~,~,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

P = 100; %$100m hedge fund value
PnL = (returns./100).*P; %calculate monthly P&L ($m)

%if monthly returns -> reporting frequency f = 1 for monthly moments
f = 1;

alpha = 0.05; %significance i.e. 95% confidence level
mES = fES(PnL,f,alpha); %monthly ES ($m)

OUTPUT:

monthly ES ($m) = -2.190
```

Source 7.11 MATLAB® code for the calculation of the monthly ES at the 95% confidence level

shows that the ES is estimated at $2.190m for the hypothetical CTA Index over the period 2008 to 2013.

Although ES is a more conservative and a useful indication of the estimated VaR at a particular confidence level, it is important to note that the ES figure does not give any indication of the severity of loss by which VaR is exceeded

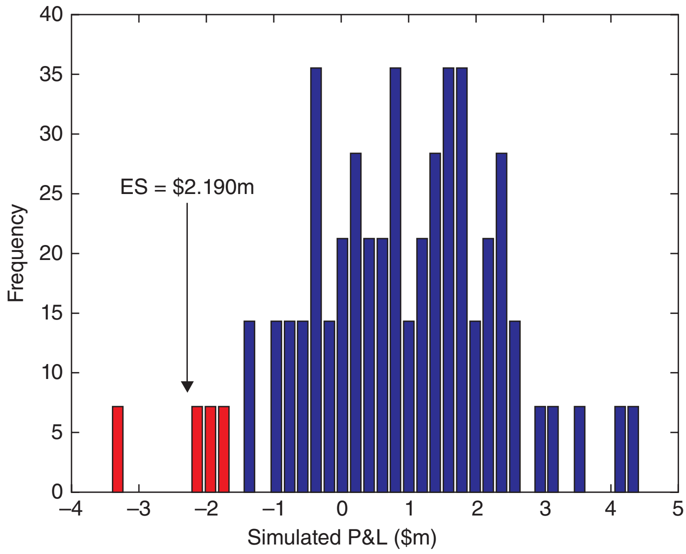


Figure 7.7 Monthly ES at 95% confidence level

```
%File: risk.m
%import XL data
[~,~,returns] = getXLData('hfma_matlab_data.xlsx','CTA Index');

AuM = 100; %$100m hedge fund value
PnL = (returns./100).*AuM; %calculate monthly P&L ($m)

%if monthly returns -> reporting frequency f = 1 for monthly moments
f = 1;

alpha = 0.05; %significance i.e. 95% confidence level

%Risk measures
mVaR = fVaR(PnL,f,alpha); %monthly VaR ($m)
mMVaR = fMVaR(PnL,f,alpha); %monthly MVaR ($m)
mES = fES(PnL,f,alpha); %monthly risk ($m)
```

Source 7.12 Monthly VaR, MVaR and ES at 95% confidence level

Table 7.2 Comparison of monthly VaR, MVaR and ES

	VaR (95%)	MVaR (95%)	ES (95%)
Monthly ($m)	−1.573	−1.636	−2.190

only the expected (or average) loss. Moreover, if considering a 99% confidence in the left-tail then, for 100 P&L values, only 1% of 100, i.e. 1 value will be used to determine the average of the tail loss. In this case, it is necessary to investigate further the area of extreme losses with a deeper analysis of the left-tail of the P&L distribution.

Table 7.2 shows a comparison of the three risk measures and clearly each method gives a more conservative estimate of monthly VaR.

7.5 EXTREME VALUE THEORY

We have already stated and discussed the fact that hedge fund return distributions have negative skewness and positive excess kurtosis, i.e. the distribution does not adequately capture the probability where losses are severe, i.e. left-tail. The thinner the tails of the normal distribution the larger will be the underestimate of the magnitude of the tail losses. This is where the branch of mathematics known as Extreme Value Theory[11] (EVT) becomes a very useful tool to apply

[11] Some hedge fund managers prefer to use *stress testing* and *scenario analysis* rather than EVT to estimate their exposures to tail events.

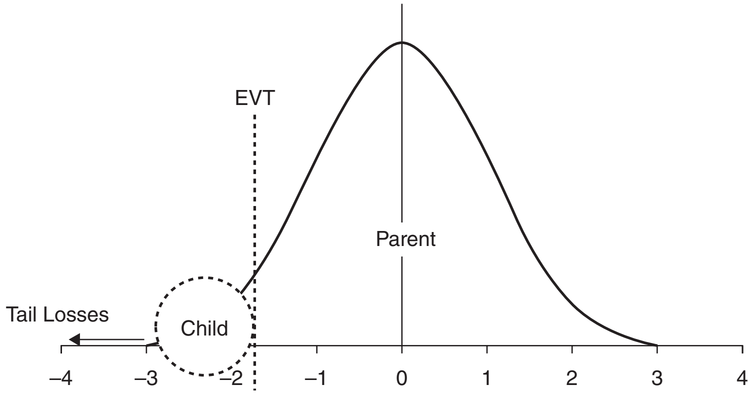


Figure 7.8 The "child" and "parent" distributions

to such a problem. The theoretical foundations of EVT were first developed heuristically by Fisher and Tippett (1928) and have since been applied to insurance and finance by Embrecht *et al* (1999). In general, EVT is the theory of modelling and measuring events which occur with very small probability. Clearly this is useful for analysing the extreme losses (i.e. left-tail) in the return (or P&L) distribution. Indeed, EVT is really the only method of extracting an accurate measure of the estimated loss given the limited data around an extreme event. Figure 7.8 shows the "parent" distribution characterised by the first two moments (i.e. μ and σ) and a separate "child" distribution that specifically characterises the distribution of losses in the left-tail of the parent distribution, i.e. the extreme tail losses. Interpreting the main distribution in terms of the parent and then a child distribution within the parent gives us some qualitative understanding of how EVT is applied.

Consider Figure 7.9 that shows a hypothetical plot of the losses for a particular hedge fund over a 100-day period. Given such information the obvious next

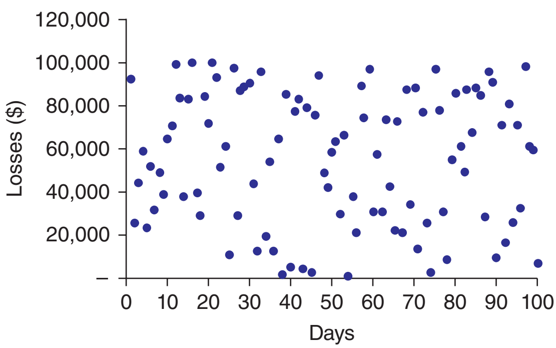


Figure 7.9 Hypothetical hedge fund losses over 100 days

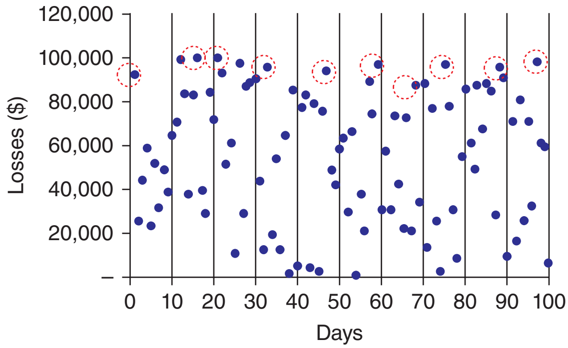


Figure 7.10 The largest losses in each 10-day block

step is to try to characterise the losses over this period in some mathematical way. EVT offers two fundamental methods for such characterisations, namely the Block Maxima and Point Over Threshold methods.[12]

7.5.1 Block Maxima

The Block Maxima (BM) method is based on subdividing the time period into a set of buckets (or blocks) of equal size. For example, Figure 7.10 reproduces the loss data given in Figure 7.9 for the hypothetical hedge fund over the 100-day period but divides the time into 10 blocks, i.e. each block is a 10-day period that will contain a certain number of losses. Taking the maximum loss in each of the 10-day blocks gives us 10 local block maxima, i.e. 10 data points which can be used to characterise (or fit) a probability distribution. This is often known as a generalised extreme value (GED) distribution.

Within the hedge fund arena, the Block Maxima method is the less preferred estimation of VaR using the EVT.

7.5.2 Peaks Over Threshold

The Peaks Over Threshold (POT) method is a more modern and widely accepted method for estimating VaR although mathematically somewhat demanding. The basic idea behind the POT method is to choose a numerical threshold to which every loss over that threshold is considered an extreme loss. The number of data points over the threshold can be used to characterise (or fit) a probability distribution, i.e. in this case it is known as the generalised Pareto[13] distribution

[12] Both methods are mathematically challenging and beyond the scope of the book and the reader is referred to cited references for more detailed explanations.

[13] Vilfredo Pareto (1848–1923) was an Italian engineer, sociologist, economist and philosopher and made several important contributions to economics, particularly in the study of income distribution.

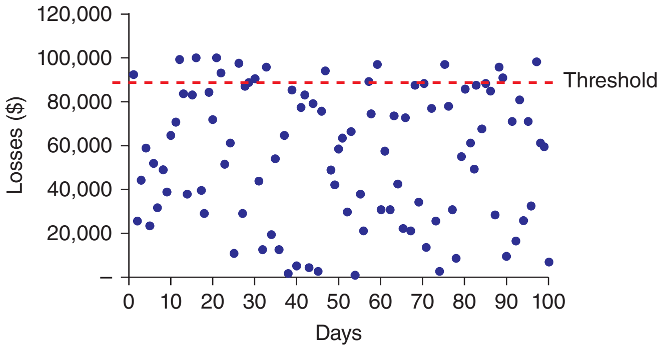


Figure 7.11 Losses over a threshold

(GPD). Figure 7.11 shows an example of a typical threshold in which only those losses exceeding \$90,000 will be considered extreme relative to the value of the portfolio.

We have already seen in Chapter 3 how to consider a probability distribution in terms of a CDF, i.e.:

$$F(x) = P(X \leq x) \tag{7.10}$$

i.e. for a given value, x, $F(x)$ is the probability that observed value of a random variable X will be at most x, i.e. less than or equal to x. For EVT there is a different function to consider based on a conditional probability (in a similar way to how the conditional expectation models the ES). That is, for an excess distribution of the return variable X over a certain threshold u, the conditional probability distribution of $y = X - u$ such that $X > u$ can be written as:

$$F_u(y) = P(X - u \leq y \,|X > u) \tag{7.11}$$

Figure 7.12 shows a schematic representation of the excess distribution of the return variable X over a certain threshold u.

The CDF is now the probability that $y = X - u$, i.e. the excess loss (or exceedance) over the threshold is less than or equal to y conditional on X exceeding the threshold u. One of the main theorems of EVT developed by Pickands (1975) states that for a reasonably high threshold, u, $F_u(y)$ can be approximated by the General Pareto Distribution (GPD) which can be written as:

$$G(X) = \begin{cases} 1 - \left(1 + \dfrac{\xi y}{\beta}\right)^{-1/\xi} & \textit{if } \xi \neq 0 \\ 1 - \exp\left(-\dfrac{y}{\beta}\right) & \textit{if } \xi = 0 \end{cases} \tag{7.12}$$

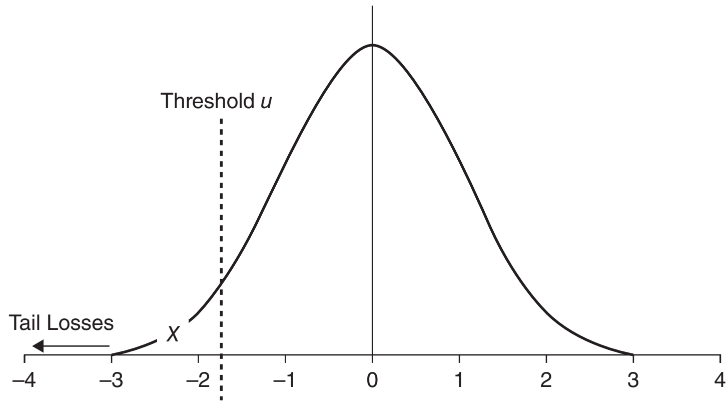


Figure 7.12 The random variable X and threshold u

where $y = X - u$, $\xi = 1/\alpha$ is the shape parameter, α the tail index and β a simple scaling parameter.

There are several approaches to estimating the parameters in Equation 7.12, such as maximum likelihood, elemental percentile method and the method of moments described in the studies by Hosking and Wallis (1987), Grimshaw (1993) and Castillo and Hadi (1997).

An estimate of the VaR using the GPD approach described above can be written as:

$$VaR_{1-\alpha} = u + \frac{\beta}{\xi}\left(\left(\frac{N}{n_u}\alpha\right)^{-\xi} - 1\right) \tag{7.13}$$

where N is the total number of data points and n_u the number of data points that exceed the threshold u.

Furthermore, the method can be extended so that the expected shortfall $ES_{1-\alpha}$ can stated in terms of the $VaR_{1-\alpha}$, that is:

$$ES_{1-\alpha} = \frac{VaR_{1-\alpha}}{1-\xi}\frac{\beta - \xi u}{1-\xi} \tag{7.14}$$

For hedge funds, the amount of data in the tail of the returns distribution is often small and therefore leads to broad confidence intervals and weak significance estimates. Both the BM and POT method suffer from the problem of limited data although it is possible to reduce the time division in the BM method or lower the threshold for the POT technique to produce more data points to fit to the desired distribution.

In this chapter we have provided an introduction to the main quantitative risk measures from the traditional VaR approaches to some of the more advanced and challenging theoretical hedge fund market risk models. We have seen throughout the book that hedge fund returns usually have fatter tails and an asymmetric return distribution which clearly violates the assumption of a normal distribution that underlies traditional measures. In order to address such limitations many extensions to the traditional VaR methods have been developed as better estimates of hedge fund market risk. Such methods either explicitly incorporate skewness and kurtosis into the model or focus primarily on the left-tails of the return distribution where most of the large losses occur. Despite the availability of more robust and potentially accurate market risk models, it must be pointed out that the analysis here covers one component of the risks associated with hedge funds. A more complete treatment would involve incorporating other equally important risks into the analysis. For example, the monitoring, management and reporting of credit, liquidity and operational risk should also be considered alongside market risk within a robust and effective hedge fund risk management process.

References

Amenc, N. and Martellini, L. (2003b), *The Brave New Work of Hedge Fund Indices*, EDHEC Risk Institute Publications.

Amenc, N., Sfeir, D. and Martellini, L. (2003a), "An Integrated Framework for Style Analysis and Performance Measurement", *The Journal of Performance Measurement*, Summer, 35–41.

Anson, M. (2008), "The Beta Continuum: From Classic Beta to Bulk Beta", *The Journal of Portfolio Management*, Winter.

Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1999), "Coherent Measures of Risk", *Mathematical Finance*, 3(9) 203–228.

Bailey, J.V. (1990), "Some Thoughts on Performance-Based Fees", *Journal of Portfolio Management*, 4, July–August, 31–40.

CAIA Level 1, (2009), Volume 1, *Study Handbook*, Uppermark.

Carhart, M.M. (1997), "On Persistence in Mutual Fund Performance", *Journal of Finance*, 52(1), March, 57–82.

Castillo, E. and Hadi, A. (1997), "Fitting the Generalised Pareto Distribution to Data", *Journal of the American Statistical Association"*, 92(440) 1609–1620.

Cornish, E. and Fisher, R. (1937), "Moments and Cumulants in the Specification of Distributions", *Review of the International Statistical Institute*, 307–320.

Darbyshire, P. and Hampton, D. (2011), *Hedge Fund Modelling and Analysis Using Excel and VBA*, Wiley Finance.

Embrecht, P., Klüppelberg, C. and Mikosch, T. (1999), *Modelling Extremal Events*, Springer-Verlag, Berlin.

Fama, E. and French, K.R. (1992), "The Cross-section of Expected Stock Returns", *Journal of Finance*, 47(2), June, 427–465.

Favre, L. and Ranaldo A. (2005), "Hedge Fund Performance and Higher-Moment Market Models", *The Journal of Alternative Investments*, 8(3) 37–51.

Fisher, R.A. and Tippett, L.H.C. (1928), "Limiting Forms of the Frequency Distributions of the Largest or Smallest Member of a Sample", *Proceedings of the Cambridge Philosophical Society*, 24, 180–190.

Fung, W. and Hsieh, D.A. (2004), "Hedge Fund Benchmarks: A Risk-Based Approach", *Financial Analysts Journal*, 60(5) 65–80.

Géhin, W. (2007), "The Challenge of Hedge Fund Performance Measurement: a Toolbox Rather Than a Pandora's Box", *EDHEC Risk Institute*, January.

Goodwin, T.H. (1998), "The Information Ratio", *Financial Analysts Journal*, 54(4), July–August, 34–43.

Graham, J.R. and Harvey, C.R. (1997), "Grading the Performance of Market Timing Newsletters", *Financial Analysts Journal*, November–December, 54–56.

Gregoriou, G.N. and Gueyie, J.P. (2003), "Risk-Adjusted Performance of Funds of Hedge Funds Using a Modified Sharpe Ratio", *Journal of Wealth Management*, 6(3), Winter, 77–83.

Grimshaw, S.D. (1993), "Computing Maximum Likelihood Estimates for the Generalised Pareto Distribution", *Technometrics*, 35, 185–191.

Gupta, R.H., Kazemi, H. and Schneeweis, T. (2003), "Omega as a Performance Measure", *CISDM Research Paper*, June.

Hampton, D.E. (2009), "A Mean-Variance Capital Asset Pricing Model for Long Short Equity Portfolios", March, SSRN Working Paper http://ssrn.com/author=837926.

Hosking, J.R.M. and Wallis, J.R. (1987), "Parameter and Quantile Estimation for the Generalised Pareto Distribution", *Technometrics*, 29(3) 339–349.

Jarque, C.M. and Bera, A.K. (1987), "A Test for Normality of Observations and Regression Residuals", *International Statistical Review*, 55(2) 163–172.

Jensen, M.C. (1968), "The Performance of Mutual Funds in the Period 1945–1964", *Journal of Finance*, 23(2), May, 389–416.

Kaplan, P.D. and Knowles, J.A. (2004), "Kappa, a Generalized Downside Risk-Adjusted Performance Measure", *Journal of Performance Measurement*, 8(3) 42–54.

Kat, H. and Miffre, J. (2002), "Performance Evaluation and Conditioning Information: the Case of Hedge Funds", Working Paper, University of Reading.

Keating, C. and Shadwick, W. (2002), "A Universal Performance Measure", *Journal of Performance Measurement*, 6(3) 59–84.

Lo, A.W. (2002), "The Statistics of Sharpe Ratio", *Financial Analyst Journal*, 58(4), July–August, 36–52.

Lhabitant, F.-S. (2004), *Hedge Funds – Quantitative Insights*, Wiley Finance.

Lhabitant, F.-S. (2007), *Handbook of Hedge Funds*, Wiley Finance.

Malkiel, B.G. (1995), "Returns from Investing in Equity Mutual Funds 1971 to 1991", *Journal of Finance*, 50(2) 549–572.

Mandelbrot, B. (1982), *The Fractal Geometry of Nature*, New York, WH Freeman.

Mantegna, R.N. and Stanley, H.E. (2000), *An Introduction to Econophysics – Correlations and Complexity in Finance*, Cambridge University Press.

Markowitz, H. (1952), "Portfolio selection", *Journal of Finance*, 7, 77–91.

MATLAB® (2013) *Statistics Toolbox User Guide*, Mathworks.

MATLAB® (2013) *Financial Toolbox User Guide*, Mathworks.

MATLAB® (2013) *Optimisation Toolbox User Guide*, Mathworks.

Merton, R.C. (1990), *Continuous-Time Finance – A Dynamic General Equilibrium Model of the Asset Market*, Blackwell.

Modigliani, L. (1997), "Are Hedge Funds Worth the Risk?", U.S. Investment Perspectives, Morgan Stanley, Dean Witter, December.

Modigliani, F. and Modigliani, L. (1997), "Risk-Adjusted Performance", *Journal of Portfolio Management*, 23, 45–54.

Pickands, J. (1975), "Statistical Inference using Extreme Order Statistics", *Annals of Statistics*, 3, 119–131.

Rockafellar, R.T. and Uryasev, S. (2000), "Optimisation of Conditional Value-at-Risk", *Journal of Risk*, 2(3) 21–41.

Sharma, M. (2004), "A.I.R.A.P. – Alternative Risk-Adjusted Performance Measures for Alternative Investments", *Journal of Investment Management*, 2(4) 34–65.

Sharpe, W. (1994), "The Sharpe Ratio", *Journal of Portfolio Management*, Fall, 49–58.

Sortino, F., van der Meer, R. and Plantinga, A. (1999a), "The Dutch triangle: a framework to measure upside potential relative to downside risk", *Journal of Portfolio Management*, Fall, 50–58.

Sortino, F., van der Meer, R. and Plantinga, A. (1999b), "The Upside Potential Ratio", *Journal of Performance Measurement*, Fall, 4(1) 10–15.

Stutzer, M. (2000), "Portfolio Performance Index", *Financial Analysts Journal*, 56, May–June, 52–61.

Treynor, J.L. (1952), "Toward a Theory of Market Value of Risky Assets", Unpublished manuscript printed as Chapter 2 in *Asset Pricing and Portfolio Performance*. Edited by Robert A. Korajczyk (1991), *Risk Books*.

Xu, X.E., Liu, J. and Loviscek, A.L. (2009), "Hedge Fund Attrition, Survivorship Bias, and Performance: Perspectives from the Global Financial Crisis", Working Paper, Seton Hall University.

Yamai, Y. and Yoshiba, T. (2002), "On the Validity of Value-at-Risk: Comparative Analyses with Expected Shortfall", *Monetary and Economic Studies*, January.

Index