UNIVERSITY OF SOUTHAMPTON

Constructing Smart Financial Portfolios from Data Driven Quantitative Investment Models

by

Chetan Saran Mehra

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the
Faculty of Physical and Applied Sciences
Electronics and Computer Science

November 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES
ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Chetan Saran Mehra

Portfolio managers have access to large amounts of financial time series data, which is rich in structure and information. Such structure, at varying time horizons and frequencies, exhibits different characteristics, such as momentum and mean reversion to mention two. The key challenge in building a smart portfolio is to first, identify and model the relevant data regimes operating at different time frames and then convert them into an investment model targeting each regime separately. Regimes in financial time series can change over a period of time, i.e. they are heterogeneous. This has implications for a model, as it may stop being profitable once the regime it is targeting has stopped or evolved into another one over a period of time. Changing regimes or those evolving into other regimes is one of the key reasons why we should have several independent models targeting relevant regimes at a particular point in time.

In this thesis we present a smart portfolio management approach that advances existing methods and one that beats the Sharpe ratio of other methods, including the efficient frontier. Our smart portfolio is a two-tier framework. In the first tier we build four quantitative investment models, with each model targeting a pattern at different time horizon. We build two market neutral models using the pairs methodology and the other two models use the momentum approach in the equity market. In the second tier we build a set of meta models that allocate capital to tier one, using Kelly Criterion, to build a meta portfolio of quantitative investment models. Our approach is smart at several levels. Firstly, we target patterns that occur in financial data at different time horizons and create high probability investment models. Hence we make better use of data. Secondly, we calculate the optimal bet size using Kelly at each time step to maximise returns. Finally we avoid making investments in loss making models and hence make smarter allocation of capital.

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

EW = Equally weighted.
MVO = Mean Variance Optimisation.
FK = Fractional Kelly.
OK = Optimal Kelly.
MK = Median Kelly.
K KF = Kelly with Kalman Filter.
MK KF = Median Kelly with Kalman Filter.
K MA = Kelly with Moving Average.
MK MA = Median Kelly with Moving Average.
OU = Ornstein Uhlenbeck model.
QIM = Quantitative Investment Model.
OLS = Ordinary Least Squares.
VR = Variance Ratio.
VRP = Variance Ratio Profile.

Declaration of Authorship

I, Chetan Saran Mehra, declare that the thesis titled "Constructing smart financial portfolios from data driven quantitative investment models" and the work presented in the thesis are both my own, and have been generated by me as the result of my original research. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have specified exactly what was done by others and what I have contributed myself;
- Part of this work has been published in a paper presented at a conference (see Section 1.4). None of this work has been published before submission.

Chetan Saran Mehra ...

Acknowledgements

My journey during this PhD was an arduous one and I am grateful to a number of people for their help. Firstly I would like to thank my supervisors Dr. Adam Prügel-Bennett and Dr. Enrico Gerding, for being fantastic guides and friends to me throughout my research. Enrico and Nick Jennings taught me how to write properly. I am particularly indebted to Adam for taking me under his wing and propping me up when I was down and out. I will never forget the calming effect of Adam's words on me. I always walked out of his office having learnt something, feeling happier and full of positive energy. I am also indebted to Professor Mahesan Niranjan, who took time out to explain concepts to me even when he didn't need to. I will miss all three of you very much and I hope I can be a tenth as insightful and accomplished as the three of you.

I am also indebted to Dr. Valentin Robu who was a wonderful friend, colleague and a dinner partner late at night in the lab. Valentin was a constant source of nuts and biscuits, he helped me refine the language for my paper and in general always gave me sound advice. Long was also a great help and a perfect partner to discuss ideas especially from a mathematical perspective. I am also deeply indebted to my friends in the lab, especially Lampros and Beining for being there for me when I just needed to talk, get my code or equations checked not to mention all the proof reading and feedback regarding my work.

I am thankful for the friendship and support of Lampros, Thanos, Bening, Matteo, Long, Henry, James, Moody, Meritxell, Seb, Rama, Rueben, Sasha and David. I am also grateful to Jonathan Hare and Jules for sharing their music without which I would have gone insane. Finally I am grateful to Dr. Eric Cooke for getting me on track with my work and putting in a good word to Adam.

Chapter 1

Introduction

Financial markets are a large, rich and continuous source of data. This data plays a very important role in our lives both directly and indirectly. In our current economic system, we are dependent on financial markets and financial securities for numerous everyday needs, such as insurance, mortgages, automobile loans, financial savings as well as old age pensions. Hence, it is important that we take informed decisions, whether these relate to the size of our loans or the management and investment of our life savings. When it comes to life savings and old age pensions global financial markets have played a very important role. Today, whether we can afford to send our children to university or afford good care in our sunset years, depends not only on how much we save, but crucially on how well we manage our savings. With an aging population that is living longer, the ability to fund their pensions for longer periods means that governments, life insurance companies and asset managers are facing a very real challenge (Banks et al., 2002).

Most of the returns for asset managers are generated through investments made in global financial markets (Timmermann et al., 1999). Asset managers are dependent on strong positive performance from these markets. They suffer immensely when these markets do not perform well or worse when markets are negative. In order to overcome this problem, asset managers look for investment opportunities and investment methods that are not entirely dependent on sustained positive performance of these global markets by investing in illiquid investments that are not readily traded, such as real estate and private equity. However, these types of illiquid investments usually have long lock-in periods, making efficient reallocation of capital difficult.

Managing investments in liquid markets is relatively simple but it comes with its own challenges. Since the mid 1980s to the year 2000, major world markets were in a strong upward trend, popularly referred to as a bull market (Figure 1.1). That uptrend made investing easy, as in all bull markets, nearly all stocks went up in price. Most fund managers would only worry about relative underperformance to the market or the main index such as FTSE 100 or S&P 500. When the internet bubble burst markets crashed taking the world into a recession, hurting investments and pensions globally. Most UK company pensions were in negative equity and at one stage the estimated shortfall was close to £800 billion. Since the internet bubble burst in March 2000, global markets have lurched from one crisis to another including the financial crisis of 2008 (Figure 1.1). This continued state of flux ensured that global markets did not cross their previous peak. Between 2000 to 2012 global markets struggled, not making new highs hurting the pensions and life savings of many.

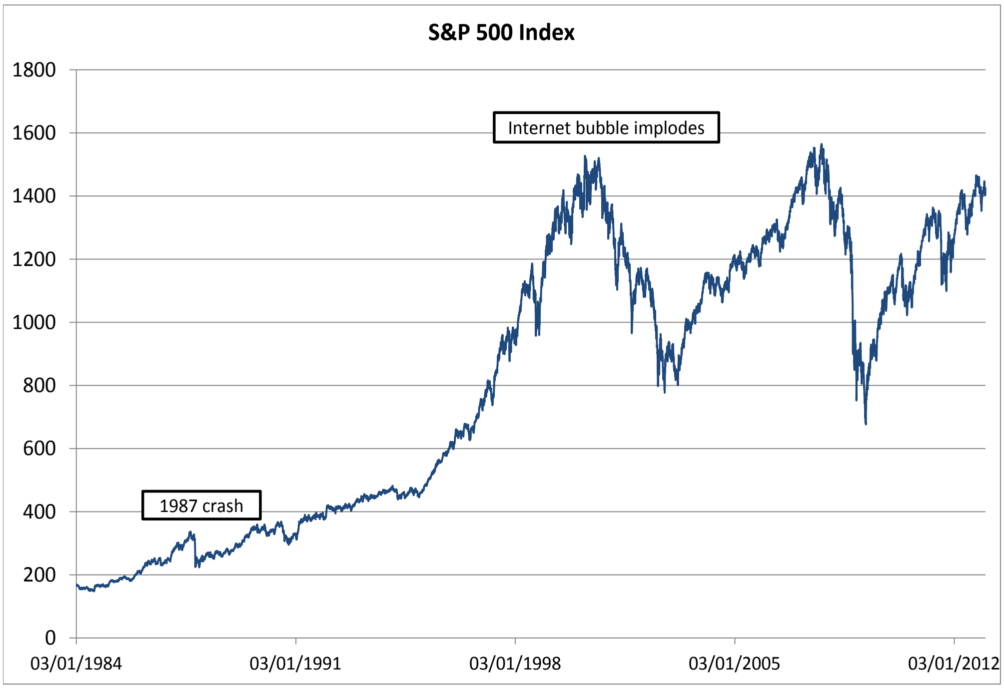


FIGURE 1.1: Specimen of S&P 500 Index price (1984 to 2012). Here we see a tremendeous positive trend in price from 1984 till the internet bubble burst in the year 2000. Since then the market has struggled to follow its previous secular positive trend hurting long term investors, and attaining new highs only in 2013.

With pension funds not generating enough profits owing to markets that have not performed well, in the last ten to twelve years a niche area of the asset management industry has started to flourish, namely hedge funds. A hedge fund is essentially an asset manager that specialises in a rather arcane but sophisticated manner of managing money. Hedge funds have flexible mandates that give them more room to be innovative in their methods. Although the hedge fund industry is at least 60 years old, it was always small as it offered some very specific benefits, such as returns that were independent or uncorrelated to the broad market. Observing the consistent performance of hedge funds, large pension fund trustees have started to give them more capital to invest. News media such as the Financial Times estimates their current size to be close to GBP 2 trillion.

Within hedge funds there is a subset of portfolio managers who manage their portfolios through quantitative investment models. These hedge fund managers may have investment methodologies that are similar in approach to other portfolio managers in the hedge fund arena. However they are different in that their investment decision making is driven by quantitative models. More specifically quantitative portfolio managers work at the intersection of finance, mathematics and computing. They strive to find ways to identify anomalies or statistically important and stable relationships in data, which can be modelled and used to build a portfolio of securities.

Historically quantitative investment models for portfolio management were initially pursued in the 1960s by the investment management arm of Wells Fargo bank which hired academics to start tracking major stock indices such as S&P 500 based on the observation that most portfolio managers underperformed the market. Wells Fargo sold its business to Barclays Bank which renamed it Barclays Global Investors (BGI) which eventually became the largest asset manager in the world. BGI got bought by BlackRock which is now an even larger asset manager. However, most of the assets at BlackRock simply track or replicate broad stock market indices.

Finance as a whole became rapidly quantitative in 1952 with the publication of Harry Markowitz's mean variance framework for portfolio construction (Markowitz, 1952), the multi-factor portfolio model by Ross (1976) and in 1973 with the publication of the Black and Scholes (1973) option pricing formula. With these breakthroughs the market and its participants have continued to adapt and became more quantitative and today are nearly totally driven by computers, especially in the case of exchange listed securities (Mackenzie, 2008).

In recent years quantitative portfolio managers have been borrowing mathematical models from other domains, such as signal processing and machine learning with the expectation to build better models for markets. In the last few years models for real time high frequency data have been the main battleground for practitioners to build cutting edge mathematical tools combined with low latency connectivity and powerful computers. Some argue that this is a natural evolution while others think it is harming markets. Whatever the case may be about these models there are three universally applicable components to portfolio construction be it in high frequency or low frequency investment. The three components to managing a portfolios of securities are:

- High probability investment opportunity (signal),
- Correct investment size (investment capital allocation),
- Managing portfolio risk (investment capital allocation).

These three components are key to good portfolio management and they are presented in order of importance. At the outset it is important to have an investment opportunity that has a high probability of success (much better than 50:50). Once this high probability investment opportunity is identified, one needs to determine how much capital to invest. The amount of capital invested directly leads to how much risk is taken, since the amount of capital invested determines the exposure to an investment and its related variance. In our thesis we will focus on creating high probability investment opportunities thorough quantitative investment models and the correct investment size for these investment opportunities.

1.1 Problem Statement

Quantitative portfolio managers deal with diverse and rich sources of data in their work. The data can take the form of prices, earnings, debt, yields as well as ratios such as price to earnings, book to price, debt to equity, etc. However, financial data is noisy and it is a challenge to extract useful structure or information from such data. Useful information or patterns or structure in data can exist in different time horizons and, large data sets can create further patterns while interacting with each other. For example a pattern may exist in a 10 period window but may appear at irregular intervals. Further, a pattern may exist at a 200 period window, just like the 10 period window, while no discernible pattern may exist in a 20, 30 or 40 period window. This means that one could have two different processes in the data and it is possible to build a mathematical model for 10 and 200 period windows of data (Figure 1.2).

FIGURE 1.2: Specimen of Software AG's price. on very short horizons of approximately 20 to 30 days one can see mean reverting behaviour, we see some momentum on longer horizons of 100 days and then mean reversion again on life of the data set of three years.

Broadly there are two main categorisations within which markets are found to operate namely momentum and mean reversion. Momentum is the tendency of investments to exhibit persistence in their relative performance (Moskowitz et al., 2011). That is, investments that have performed relatively well in the past, continue to perform relatively well in the future; and those that have performed relatively poorly, continue to perform relatively poorly. A range bound or mean reverting market is the opposite of momentum, i.e. the tendency of investments to exhibit no persistence in their relative performance and generally appear not to move up or down for any extended period of time. However, financial data is noisy and mean reversion occurs even in momentum driven markets, i.e. both processes are present and evolving at the same time but usually identified over different time horizons. The key challenge is that capturing mean reversion and momentum successfully requires models that are dedicated to each of these processes.

In practice, most fund managers focus either on long run momentum models or for range bound markets, short run mean reversion models. Models for long run momentum and short run mean reversion have very different risk-return profiles and because they cater to a niche investor, it is simpler for an investor to categorise them as one or the other. That said a single model is unlikely to perform well in all market conditions, i.e. momentum and range bound markets owing to different regimes and their nature. Given this challenge portfolio managers should have several quantitative investment models that generate portfolios.

Combining quantitative investment models to form a portfolio has the benefit of using the data in a comprehensive manner. By capturing specific patterns found in market data at different time horizons, it is possible to benefit from information extraction from the market in a thorough and complete way. Diversified portfolios generally tend to be more stable and resilient by reaping benefits in volatile conditions as well as in changing market conditions.

1.2 Research Question

As discussed above there are differing processes in financial data at different time horizons and the best way to capture them is to have dedicated models targeting these processes. For our research the key research question is can a portfolio of quantitative investment models provide superior risk adjusted returns when compared to a portfolio based on a single quantitative investment model. This key research question also depends on our ability to identify and model momentum and mean reversion in financial data.

Although there has been some research work on constructing portfolio of quantitative investment models; there is limited published research on constructing a portfolio of quantitative investment models that are dedicated to patterns found in data at different time horizons, as well patterns in data, that are generated from interactions within large multivariate data sets. Given this we need a principled approach for studying the data and building models that suit and work well with that data set.

1.3 Research Challenges

Quantitative portfolio managers are notoriously secretive about their methods and tools. They keep their technology and intellectual property under wraps since their success depends on their ability to have an advantage over their competitors and other investors in general. Therefore there is limited academic literature published on this problem, as well as very little discussion about methods and methodologies used by these portfolio managers. However, some academics who either advised or operated a hedge fund such as Jean-Philippe Bouchaud of Capital Fund Management and Edward Thorp of Newport Partners who ran one of the most successful quantitative hedge funds in the world, have published their research about methods and tools used in this arena (Thorp, 1967).

Another structural challenge is the inability to capture all strategies in the quantitative hedge fund space, even those that are quite suitable for quantitative modelling. Specifically exchanges have transparent pricing but sometimes certain instruments are synthetically created by banks for example, options and warrants, their prices are not easily available and the value is not determined in a transparent manner. In this context strategies that use asset classes which have exchange traded instruments such as equities, index futures, government bonds, foreign exchange (FX), (although FX trades both off and on exchanges these days) are best suited to quantitative modelling as they offer:

1. Prices that are available to the public from stock exchanges,
2. Clean and consistent historical data that is available to the public to build and test models with and,
3. Open competitive stock exchanges in a well regulated environment.

Quantitative investment models for an asset class where prices are not transparent and price history is not available, (e.g. where prices are negotiated over the telephone or are generally off the exchange) are not suitable even though mathematically they would make good candidates for quantitative strategies. Another key challenge is building statistical models, considering financial data is noisy and heterogeneous. Given these challenges we will build four quantitative investment models that will capture structure in the data. We will then build a meta model that will allocate capital to these quantitative investment models, that will give us a portfolio of quantitative investment models.

1.4 Research Contributions

Given the limited body of academic literature with reference to this subject, we will contribute to the body of research literature for portfolio construction when it comes to combining quantitative investment models primarily through our proposed meta models. Specifically, our key research contribution will be our meta models that present a novel way of combining four quantitative investment models that focus on equities as an asset class at different time horizons. We use stocks and index futures data to build two market neutral models that are most suitable for range bound markets. Our third model is a momentum model that is suitable for momentum driven markets and the fourth model is a constant rebalanced portfolio models for stocks. To address these questions we use both simulated as well as real financial time series data which is distributed by exchanges daily.

Our initial findings and novel model were published in Mehra et al. (2014). This paper presents our meta models, which demonstrated better risk adjusted performance for two years of out-of-sample data when compared to existing methods. Furthermore our meta model uses the Kelly criterion to build a portfolio of quantitative investment models. We systematically test several versions of Kelly criterion using both synthetic and real financial data.

1.5 Thesis Outline

The rest of the thesis is written in eight chapters. This section presents the structure of the thesis and a brief outline of each chapter.

Chapter 2: Background & Related Work

This chapter is devoted to a literature review where we focus on the history of portfolio construction, new developments that have influenced the field, as well as the challenges we face in portfolio construction especially when we have noisy and non-stationary financial data. We begin with a discussion on financial time series and the properties found in many financial time series data. Subsequently, we discuss the foundations of portfolio construction, key contributions to the field, benefits and shortcoming of the primary models and how academics and practitioners have addressed these issues. We also discuss influences from outside the field of finance and how several methods from probability and information theory have influenced the field of finance and portfolio construction as a whole.

We then follow this up with a discussion on quantitative investment models used by practitioners and academics. Although there are various quantitative models such as those for option pricing, convertible arbitrage etc., we focus on models that are related to our research, i.e. ones used to build portfolios. We will primarily look at market neutral models, momentum models, etc. using financial time series data (price). We also discuss other efforts to construct a portfolio of quantitative models by researchers and practitioners.

Chapter 3: Statistical Analysis of Data to Identify Structure and Patterns in Data

This chapter presents a number of statistical tests that we use to identify relevant patterns in data for our models. We divide the statistical tests according to the type of quantitative investment model, since each of our quantitative investment models is targeting a particular pattern found in financial data at different time horizons.

Chapter 4: A Framework of Quantitative Investment Models

In this chapter we present the framework within which our models will operate. This framework will have two key tiers of quantitative investment models, one tier interacting with the market and the second tier, which is our main contribution the meta models is used for allocating capital to quantitative investment models, resulting in the meta portfolio. We will discuss how the models operate, i.e. how they interact with each other and the market as well as what the inputs and outputs are at each stage. This chapter lays out the broad map for our work. In the chapters that follow we develop the quantitative models that will populate this framework.

Chapter 5: Constructing Quantitative Investment Models

In this chapter we present the four quantitative investment models that form the first tier of the framework. We build these quantitative investment models to capture structure in data at different time horizons namely mean reversion and momentum. We build two market neutral pair models, one for global equity index futures and one for the European equity market. We also build a momentum model for global equity index futures and a long only equity model focused on the FTSE 100 constituents. We will discuss the rationale, approach and structure of these models as well as show the process and steps we took to build these models and present their respective algorithms. These models are then used in Chapter 7 on synthetic data and in Chapter 8 on real financial data, where we analyse their performance statistics. The output of these models is then used by the meta models that will allocate investment capital to them and form the meta portfolio.

Chapter 6: Constructing Meta Models

In this chapter we present our meta models that form the second tier of the framework. We have four benchmarks and five models. These five models show the evolution of our meta models. We discuss them in detail and present the algorithms and how they operate. The meta models use the output of the quantitative investment models built in the previous chapter, forming the meta portfolio.

Chapter 7: Generating Synthetic Data for Models

In this chapter we generate some synthetic data so that we can test our quantitative investment models and our meta models. The main objective of generating synthetic data is to infuse features found in prices of real financial data and to see if our models in both tiers of the framework will generate the returns and have the performance that we expect them to have. Here we will assess the performance of the quantitative investment models as well as the meta models on the basis of the Sharpe ratio, which is the standard and most widely used metric.

Chapter 8: Performance and Analysis of Models Using Real Data

In this chapter we use real financial data to test our models. We begin by presenting the data that we will use and discuss what precautions we take to ensure that the data is clean and consistent. We then divide the data into in-sample and out-of-sample periods. We use the in-sample period to calibrate our models and the out-of-sample to validate them. As in Chapter 5, we begin with our quantitative investment models and discuss the performance of the models both in the in-sample period and out-of-sample period. We then move to our main contribution, the meta model, where we show the performance of the in-sample, out-of-sample period and the full sample. We discuss the performance in detail and highlight some of the key points which make our model better and different.

Chapter 9: Conclusion & Future Work

We present a brief summary of our research, some thoughts and observation on our research. We also discuss the strengths and weaknesses of our model and suggest potential future areas of improvements to this research. We then discuss potential applications of the models to other areas of work.

1.6 Summary

In this chapter we gave a brief background to our research to give context to the reader. We also discussed our research question, challenges related to solving the problem and the research contribution. Our main contribution is a portfolio construction approach that accounts for risk while allocating capital in quantitative investment models. We also systematically test Kelly based portfolios for investing.

We also present the outline of this thesis, describing what each chapter contains. This chapter sets the broader background for our research, as well as the three questions that we asked about how a good portfolio manager breaks down the challenge of portfolio management namely investment opportunity with high probability of success, correct investment size and risk control. The next chapter will present and discuss research literature that is related to our research.

Chapter 2

Background & Related Work

In this chapter we begin by giving an overview of features found in financial time series data to understand how it is different from assumptions of homogeneous data used in standard financial and statistical models. We will highlight some of the stylised facts found in financial time series as some of these features will help us construct models. We then discuss the foundations of portfolio construction, different methods and their shortcomings. We then present Kelly criterion a powerful method based on Information Theory, which we will use to build our models. We subsequently discuss developments in finance, investment models, portfolio management, and new methods that are slowly but steadily changing portfolio management methods. Then we look at the evolution of the asset management industry, which uses quantitative models extensively for portfolio management and subsequently we discuss quantitative investment model that we will use for modelling. Finally we take a brief look at Kalman Filter which will be used in our models.

2.1 Features of Financial Time Series Data

Financial time series can take many forms. However, in our research we focus on price time series. Prices are reported by stock exchanges, which is where stocks, bonds and other assets are listed and traded. A stock exchange reports the official opening and closing prices of listed stocks or bonds. This price is used used to value portfolios, manage and assess risk and value companies.

Financial time series (FTS) in the form of price and their daily change do not exhibit the statistically stable property of homogeneity, which is generally assumed by financial models used in portfolio construction that are themselves based on standard statistical models.

Specifically standard parametric statistical models require their data to be homogeneous and assume them to be independent and identically distributed (i.i.d). However, most financial data and especially change in stock prices do not have these properties (Cont, 2001). This is a key challenge that we will face whilst building our models, since simplistic models are likely to mis-estimate key statistics. Next we highlight some important characteristics typically found in financial returns data *(definitions can be found in Appendix A*).

- **Heterogeneity:** Heterogeneity in data exists when the data is not uniform universally. Most FTS data exhibits clustered volatility and this characteristic has been found empirically since Kendall (1953), Houthakker (1961) and Osborne (1962). Clustered volatility, characterised by autoregressive conditional heteroscedasticity (ARCH) (Engle, 1982) and generalised autoregressive conditional heteroscedasticity (GARCH) (Bollerslev, 1986) models, take into account that the residuals are not constant over time.

- **Autocorrelation in returns:** FTS exhibit autocorrelation in returns, where autocorrelation is correlation in data to its own lagged values of time series data. There have been studies that show varying amounts of autocorrelation in returns data for equity indices, stocks, mutual funds, relative performance of stocks and sectors (Lo and MacKinlay, 1988). Choice of frequency such as daily, weekly or monthly can have an impact on the results of autocorrelation tests. Data can show strong autocorrelation at very high frequency such as one minute data, no autocorrelation at daily frequency then again positive autocorrelation at monthly or weekly frequency (Lewellen, 2002).

- **Gain/loss asymmetry:** In FTS, it is observed that market indices and stock prices show large negative moves but not equally large positive moves (Cont, 2001). Typically markets trend upwards in small increments over a long period of time but falls are generally quicker and decrements are large.

- **Aggregational Gaussianity:** As the time scale Δ_t on which returns are calculated increases, their distribution looks more like a normal distribution. This means that the distribution of return is not the same on all time scales (Cont, 2001).

- **Calendar effect:** Calendar effects are cyclical anomalies in market returns based on the time of the year. There are several of these effects in financial data such as i) the January effect, where returns in the month of January were found to be larger than other months (Wachtel, 1942; Haugen and Lakonishok, 1987) and ii) the weekend effect, where returns over the weekend *(Saturday and Sunday)* were found to be lower than other days (Cont, 2001; French, 1980; Cross, 1973).

- **Volume/volatility correlation:** Trading volume is related to all measures of volatility, i.e. as volatility increases, so does trading volume. This in particular can be seen in a crisis, after company announcements of both positive and negative nature as the markets processes and reacts to new information, buying and selling stocks to reposition their portfolio (Cont, 2001).

- **Intermittency:** Returns display, at any time scale a high degree of variability. This is quantified by the presence of irregular bursts in time series of a wide variety of volatility (Cont, 2001)

- **Leverage effect:** The leverage effect refers to the observed tendency of an asset's volatility to be negatively correlated with the asset's returns. Typically, rising asset prices are accompanied by declining volatility, and vice versa. As asset prices decline, companies become mechanically more leveraged since the relative value of their debt rises relative to that of their equity. As a result, it is natural to expect that their stock becomes riskier, hence more volatile (Ait-Sahalia, 2011).

- **Absence of autocorrelation in returns:** Cont (2001) points out that linear autocorrelation of asset returns are often insignificant, except for very small intra-day scale for which micro structure comes into play. This is different from observations made by Lewellen (2002) as well as Lo and MacKinlay (1988), discussed above. However, they made these observations at much lower frequency data. Analysing financial data at different frequencies can give very different results.

Stylised facts mentioned above make interesting reading and highlight some challenges as well as some opportunities. It is important to point out that observations of the data can differ depending on a) asset class, b) time horizon, c) frequency, where an asset class is a categorisation for an investable asset such as equity, bonds, FX or commodity. Essentially, one can find different aspects of the data depending on how an analyst decides to dissect the data as well as which frequency of data they use. To the best of our knowledge there has never been a comprehensive study of FTS that has looked at all the data in all asset classes on all time horizons. Given these observations FTS poses some interesting challenges. When simplistic models are applied to this data, the results are sub-optimal especially when put through an optimiser in a naive fashion. This has clear implications for portfolio construction as it makes it harder to build robust portfolios.

2.2 Portfolio Construction

In the previous section we discussed the features of financial time series data. In this section, we consider models that use this data to build portfolios. Most cautious investors are generally sceptical of investing in one security; instead they prefer to invest in a collection of securities. Such a collection of securities is referred to as a *portfolio*. Not all investments are successful, some can lose money, hence it is always advisable to have a portfolio of investments to reduce risk of loss. In a typical portfolio of stocks, risk is measured through standard deviation of returns, a well established measure of risk. The higher the standard deviation, the higher the risk. In this context, portfolio construction can be seen as a process that helps strike a balance between risk and return. Specifically, portfolio management deals with the analysis of an investment and the theory of combining these investments into a portfolio.

The first recorded mention of portfolio allocation is from the fourth century where Rabbi Issac Bar Aha proposed an equally weighted portfolio, he proposed "a third in land, a third in merchandise and third in hand" (DeMiguel, 2009). Since then there have been some advances in the portfolio construction. The first person in recent times to give portfolio construction a well defined framework was Markowitz (1952). We will discuss how this intuitive and powerful framework has dominated the literature of portfolio construction.

2.2.1 Modern Portfolio Theory

The foundation of modern portfolio construction was laid by Markowitz (1952, 1959) and Tobin (1958). In his seminal paper in 1952, Markowitz made some key observations. He observed that returns alone were not important, but rather that a balance of return and risk was important. Risk would be measured by the variance of the returns. He also observed that when two risky assets are combined their standard deviations are not additive, provided the returns from these two assets do not posses perfect positive correlation. This meant that when a portfolio of risky assets is formed, the standard deviation of the portfolio is less than the sum of the standard deviation of its constituents.

Markowitz's framework of portfolio construction is also called the mean-variance framework. To construct a portfolio the mean-variance framework has two key inputs in an optimiser: i) expected returns, estimated using historic mean return of a stock to represent gain and ii) the covariance matrix of stocks again estimated from historic returns to represent risk. The Markowitz approach builds an *efficient frontier* as shown in Figure 2.1. Here an efficient frontier is a set of portfolios considered optimal and offer the highest expected return for a certain level of risk or the lowest risk for a given level of expected return. The optimal portfolio is measured using the Sharpe ratio, which is the annualised portfolio return minus the risk free rate divided by the annualised volatility of the portfolio returns (Equation 2.1). Here annualisation is a way of standardising the data, especially for small data sets so that comparisons are consistent. Specifically, the Markowitz model generates an efficient frontier out of several investable portfolios; these portfolios show a relationship between return and risk. The highest Sharpe Ratio represents a portfolio with the best investment opportunity.

$$Sharpe\ ratio = \frac{Portfolio\ Return - Risk\ Free\ Rate}{Portfolio\ Volatility} \quad (2.1)$$

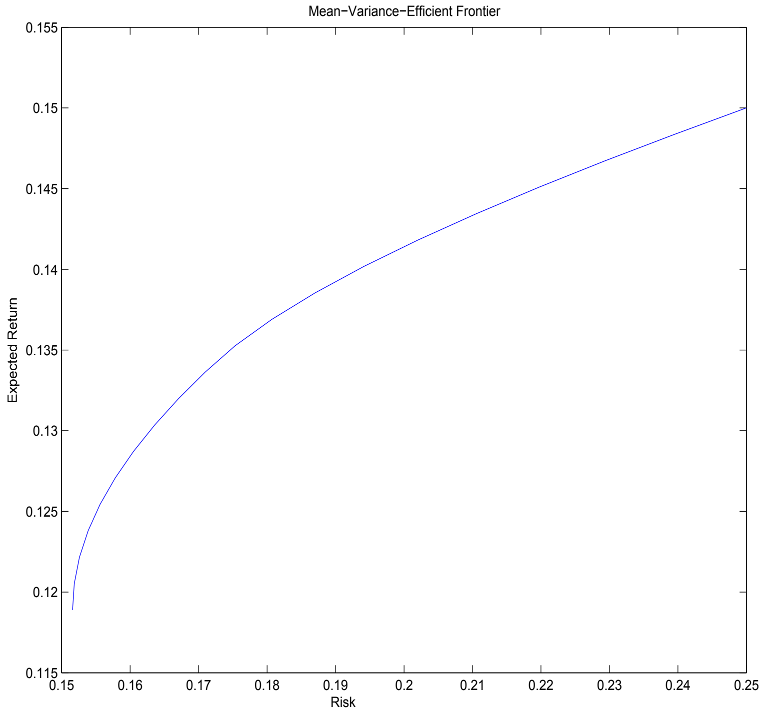


FIGURE 2.1: The efficient frontier. A specimen of the efficient frontier of three assets created in MATLAB.

Markowitz's work led to further research on portfolio theory by several researchers such as Sharpe (1964), Lintner (1965) and Black et al. (1972) who extended his work to build the Capital Asset Pricing Model (CAPM). The CAPM states, variance of a stock's price can be partly explained by the index of which that stock is a member and partly by the characteristics unique to that stock. For example variation in the price of Vodafone PLC stock can be partly explained by the FTSE 100, the main index to which it belongs. We present the CAPM Equation 2.2, which is essentially the ordinary least squares (OLS) equation.

$$R_s = \alpha + \beta R_m + \varepsilon. \quad (2.2)$$

Here, R_s is the return of a stock, α is the intercept, β is the regression coefficient that shows how much return of a stock is explained by the market index e.g. FTSE 100, denoted by R_m and ε is the residual. β is also called the Beta of stock, which measures a stock's sensitivity to the market index. High Beta means high sensitivity and *vice versa*. While analysing mutual funds Jensen (1968) showed that the Sharpe-Lintner version of the relationship between expected return and market Beta implied times-series regression exhibiting a linear relationship between the market and stocks as well as mutual funds. Merton (1973) proposed Intertemporal Capital Asset Pricing Model (ICAPM), which forecast changes in the distribution of future returns or income. The next stage of evolution was the Arbitrage Pricing Theory (APT) proposed by Ross (1976, 1980). The APT model explains the movement of a stock not explained by the index, through factors specific to a stock or further adapting it through macro economic factors such as oil, inflation and interest rates (Grinold and Kahn, 1999).

Markowitz's mean variance framework was a static one, by static we mean that estimates of returns and covariance were made from historic data and then kept unchanged. Dynamic versions of the model have been developed starting with Merton (1972), followed up by Valian (2009), Cai et al. (2013), Frey (2012), Ghosh and Mahanti (2014) and Al Halaseh and Bakar (2016). The dynamic or multi-period approach takes advantage of recent information about future expected returns. Valian (2009) shows that multi-period models perform better than single-period models in the long run. In industry some of the most sophisticated quantitative portfolio managers also practise dynamic rather than the old static approach to portfolio construction.

The mean variance framework is sensitive to data. Given the characteristics of FTS in Section 2.1, high variance in the data leads to mis-estimation of the mean, variance and covariance. Chopra and Ziemba (1993) show that using forecasts that do not accurately reflect the expected return can severally degrade performance. According to Michaud (1989) the optimising methods are sensitive to noise (variance) and end up maximising errors in estimates. Michaud also noted that small changes in expected returns (mean) changed the outcome of the portfolio dramatically. The Markowitz framework, although simple, helped formalise the problem and we shall use it as a benchmark both in our simulations as well as real financial data in Chapter 7, and Chapter 8 where we discuss our empirical results.

The model proposed by Black and Litterman (1992) called the Black-Litterman model (BL) was designed to overcome some of the shortcomings within the mean-variance framework and to combine forecasts whilst building an optimal portfolio. The BL model only differs from the Markowitz model with respect to expected returns; otherwise it is quite similar to Markowitz's mean-variance framework. The BL model starts with an equilibrium portfolio, which is essentially the market index e.g. FTSE 100, and then allows one to combine forecasts of expected returns and confidence (uncertainty) in the forecast with the equilibrium portfolio. However, for our model, we do not yet have an established equilibrium portfolio. Hence the BL methodology is not particularly useful for our objective at the moment, so we shall not include this as benchmark in our analysis, instead we will use the mean-variance model. The BL framework remains popular within the investment management industry where active investment management mandates are assessed through benchmarks.

2.2.2 Portfolio by Sorts

Another way of building portfolios is through sorting (ranking) information. This method proposed by Chriss and Almgren (2005) is based not on expected returns like Markowitz (1952) but on *ordering information* of expected returns, that is to say *ranking* of the expected returns from highest to the lowest. For example in Equation 2.3 we rank expected returns. Here r is expected return of asset $i \in (1, 2....n)$. More formally a portfolio of sorts is defined as follows.

$$r_1 \geq r_2 \geq r_3 \dots r_n. \tag{2.3}$$

Defining Sorts: *"In the most general sense a portfolio sort is a set of inequality relationships between expected returns of these assets. The simplest and most common example is a single complete sort which orders all the assets of the portfolio by expected returns from greatest to the lowest"* (Chriss and Almgren, 2005).

Sorts can be used in many ways, such as a) decile based sorts, where the top decile is the top 10 percent of a group, b) Index over/under-performers, where out-performers and under-performers of an equity index are ranked, and c) Multiple sorts, where stocks are ranked by more than one criterion.

Sorting can be done using different criteria than expected return (Asness, 1997) where we see multiple sorts used to build portfolios. Sorts are particularly useful when there are a large number of stocks to apply. The sort is used to select stocks and then uploaded into a standard mean-variance optimiser. The method is analogous to Markowitz (1952) approach as he uses information about both expected returns and covariance to build an optimal portfolio. Portfolio by sorts is very much in the spirit of the mean-variance framework. However we do not have a large number of models to use sorts in a reasonable manner.

2.2.3 Risk Parity Portfolios

Risk Parity (RP) is an asset allocation approach pioneered by Bridgewater Capital which is the world's largest hedge fund manager. Asset allocation is a higher order or top down approach to allocating capital and is related to allocating capital to an asset class. Where the main asset classes are Equities, Bonds, Foreign Exchange and Commodities. RP's unique argument is to diversify a portfolio by risk, where risk is measured through standard deviation or returns. The RP portfolio will have equal amounts of contribution to risk from each asset class. To diversify risk one generally needs to invest more capital in low risk assets than in high risk assets. The problem of extra capital is solved through leverage.

Empirical performance of RP when tested over a period of 80 years has shown to beat the classic 60/40 portfolio of stocks/bonds as shown by Asness (2012) and Maillard (2010). However, it has raised a fair number of questions about the approach as it flies in the face of CAPM and Modern Portfolio Theory. The argument that variance is the only factor that matters has also shown to be suspect in various empirical studies done by Jensen et al. (1972) and Frazzini and Pedersen (2014), where they show that high variance does not translate to high return. This not only contradicts RP but also CAPM model as well. However, the most important shortcoming of RP for us is that it cannot be implemented in its true form without the use of leverage. Owing to the use of leverage in RP we will not use this method as a benchmark to compare our model.

2.2.4 Kelly Criterion

Although Markowitz and Sharpe's work did not perform well in empirical applications (Fama and French, 2004), their work remains a *tour de force* since it helped shape the theoretical framework for the problem of portfolio construction. Meanwhile, unrelated to portfolio construction, researchers were involved in mathematical betting systems. John Kelly, a mathematician, wrote a ground breaking paper on betting (Kelly, 1956). In his paper Kelly applied concepts from information theory to a game of chance (betting). Kelly's work centred on finding the optimal "bet size", of one's capital to bet on a game of chance, that had positive expectation, such that the bettors maximised the expectation of log of wealth. Kelly's work on bet size is now known under several different terms such as *"Kelly fraction"*, *"Optimal f"*, short of optimal fraction of capital. *"Kelly bet"* and *"Kelly Criterion"*, the last term was coined by Edward Thorp. Kelly showed mathematically that his method of betting was optimal. In fact, Kelly's formula is the best way to grow capital without going bankrupt (Sinclair, 2008).

We consider making a series of n investments at time interval Δt, to grow our capital G. The gain on the investment of size X_0 is $\mu\Delta t + \sigma\Delta W$, where $E(\Delta W) = 0$ and $E(\Delta W^2) = \Delta t$. Now suppose we invest a fraction f of our capital at each time interval and put $1-f$ in a risk-free asset with return r, G represents the growth of our investment. Then our fractional return in round i is:

$$1 + (1 - f)r\Delta t + (\mu\Delta t + \sigma\Delta W_i)f. \tag{2.4}$$

And the log return is:

$$G_n(f) = \sum_{i=1}^{n} \log(1 + (1 - f)r\Delta t + (\mu\Delta t + \sigma\Delta W_i)f). \tag{2.5}$$

The expected log return is:

$$E(log(G_n(f))) = nE(\log(1 + (1 - f)r\Delta t + (\mu\Delta t + \sigma\Delta W_i)f)) \tag{2.6}$$

Expanding in Δt we get:

$$E(\log(G_n(f))) = nE((1 - f)r + f\mu)\Delta t + \sigma\Delta W_i - \frac{1}{2}\sigma^2 f^2 \Delta W_i^2 - O(\Delta t^2) \tag{2.7}$$

$$= [((1 - f)r + f\mu) - \frac{1}{2}\sigma^2 f^2]n\Delta t + O(\Delta t^2). \tag{2.8}$$

By maximising the leading order term

$$f^* = \frac{\mu - r}{\sigma^2}. \tag{2.9}$$

As shown in Figure 2.2, the horizontal axis represents Kelly fraction and the vertical axis represents growth rate of investments. As you can see in the figure, if f is too large, investors will eventually lose all of their money, even if the long run expected return is positive. If f is too small then it will take too long to make a sizeable gain, whilst at the optimal Kelly fraction f one will make optimal bets.

Kelly's method has some very useful properties. Firstly, Kelly's method of betting maximises the log of wealth (Breiman, 1961) and asymptotically maximises the geometric mean also known as the compounded rate of return of an investment that has a positive expectation. Where compounded rate of return, is the return that is earned when an investment return from a previous period stays in an investment and can begin to earn a return itself. Secondly, since Kelly is about reinvesting or is a multi-period approach it is important that an investor maximises the geometric mean. Thirdly, the expected time to reach the target wealth is minimal when using Kelly. Fourthly, the Kelly strategy is myopic, i.e. we only need to consider our current investment opportunities and wealth, not subsequent situations. Fifthly investing at a fraction (3/4, 1/2 or 1/4) of the Kelly allows investors to easily tune their desired level of risk at the expense of lower expected returns (MacLean et al., 2011). Lastly, Kelly and fractional Kelly have been shown to be optimal in Merton's continuous time finance framework by Davis and Lleo (2012), Merton (1993).

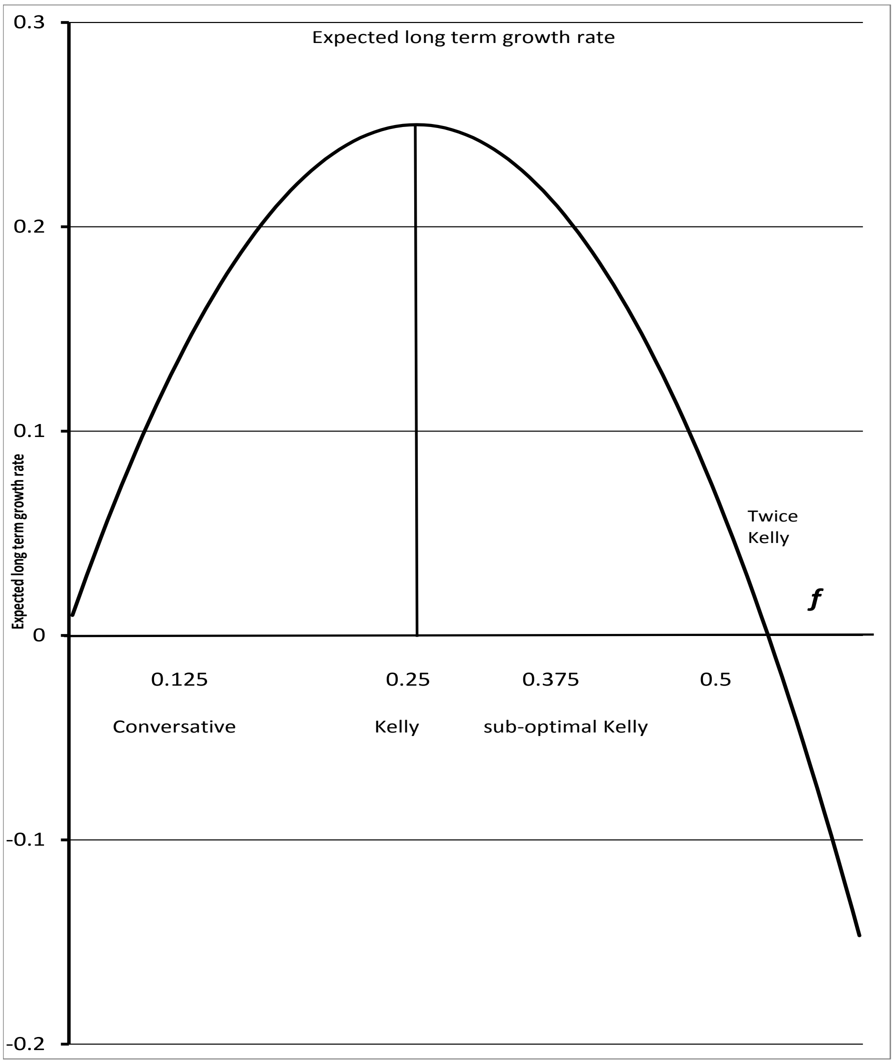


FIGURE 2.2: Expected long-term rate as a function of investment function. Specimen of a Kelly computation for binary outcomes of 50 : 50, and loss is equal to 1 the wager gain is 2 units. The highest point represents the optimal investment size 0.25, corresponding to Kelly estimates. This will maximise the growth rate of invested capital Source: (Wilmott, 2006).

Kelly's method also has some drawbacks. Firstly, a good estimate of probability of winning becomes crucial. Secondly, investment fractions get too large if probabilities are favourable, which can make bets very large as a percent of capital. Thirdly, the time necessary for long run effects to dominate other methods can be drawn out, and finally Kelly assumes that capital is infinitely divisible (MacLean et al., 2011).

Breiman was first to consider multivariate portfolio of Kelly bets allowing for overlapping sets of outcomes, called events, each with betting odds (Breiman, 1961). However Breiman did not make any suggestions how these bets should be combined, especially when the bets may exceed investable capital. To accommodate several Kelly bets fractional Kelly can be used, where fractional Kelly is a smaller proportion of the estimated Kelly bet such as 3/4, 1/2 or 1/4 of the actual Kelly fraction. MacLean et al. (2011)

conducted a study of fractional Kelly and have shown it to be optimal in Merton's continuous time framework. Rising and Wyner (2012) also presented fractional Kelly but they called it Partial Kelly. They built a fractional Kelly portfolio through a shrinkage estimator which universally shrinks Kelly estimates to manageable portfolio weights, their model essentially achieves fractional Kelly weights that one would get if one normalised Kelly estimates of several investment opportunities, giving portfolio weights. Rising and Wyners also claim that the shrinkage has the added benefit of reducing estimation error. Fractional Kelly is quite suitable in finance as often the Kelly bet size is so large that it recommends using leverage in the continuous time setting.

Maslov presented a multivariate approach of Kelly bets but it was only for uncorrelated assets (Maslov and Zhang, 1998). Laureti furthered Maslov's work for correlated assets but only in principle, using artificial data with small mean and small variance, showing that the Kelly portfolio lies on the extremes of the efficient frontier if the data is log normally distributed, i.e. its special case of the efficient frontier (Laureti et al., 2010). Laureti *et al.* state that there is hardly any difference between their approach and MVO, as they use the covariance matrix in their optimisation.

2.2.5 Universal Portfolios

The theory of Universal Portfolios (UP) was developed by Cover (1991); Cover and Ordentlich (1996). A UP is one that invests a small fraction of initial capital in a stock and the rest in a risk free deposit and rebalances the portfolio to the same weights at every time step. This is very close to the Kelly principle except that the weights are fixed. For example, if C is the capital available to invest and N is the number of stocks to invest in then the resulting allocation is based on C/N. Then at the next step one simulates all possible permutations and combinations to find the best portfolio in hindsight at $t-1$ and then use the optimal weighting from $t-1$ for the next step to give us the Best Constantly Rebalanced Portfolio (BCRP). Here the BCRP is the best outcome in hindsight.

The UP is computationally very expensive as it will do a complete grid search for all possible permutations and combinations, especially when it asymptotically achieves a log optimal portfolio, the same as Kelly. However the UP has some interesting features. For example it makes no distribution assumptions about the market, and works like a sequential or online investment algorithm. Furthermore, it is supposed to asymptotically outperform the growth rate of the best performing stock in the portfolio. Cover showed this through some well chosen examples of two stocks (Cover, 1991). However, he quickly realised that this was tougher with real data and introduced UP with side information (see Cover and Ordentlich (1996)) where the side information can be a technical indicator or correlation information, essentially bringing UP closer to our approach of having dedicated models.

Several versions of UP have been published in the last few years by Blum and Kalai (1999), Gyorfi et al. (2008), Kozat and Singer (2011) to name a few, they seem to work in theory. Most UP are rather limited case of two asset problems and asymptomatically equal to Kelly so we don't find UP useful.

2.3 Quantitative Investment Models

Quantitative models are simply ways of investing where the process is driven by an algorithm. We distinguish between quantitative models and models that are quantitative as well as systematic. By quantitative and systematic we mean that the initiation and implementation of an investment decision is entirely controlled by an algorithm, i.e. there is little to no human intervention. For our research we focus on two approaches that are quantitative and systematic in nature, a) momentum driven trend following and b) mean reversion driven market neutral. The reason we focus largely on these two features is that they i) are ubiquitous in financial data, ii) capture the most prevalent features of financial data and iii) address our goal of capturing patterns and structure in data at different time scales.

2.3.1 Momentum

Major stock market indices, commodities as well as stocks exhibited momentum or trends in their price data for at least two centuries. Momentum has been shown to be statistically significant by an in-depth study performed by Lemperiere et al. (2014). Momentum investing is also known as trend following, has a long and well established history. The earliest known momentum investors are CTAs or Commodity Trading Advisors, they initially focused on agricultural futures and cash crops. CTAs evolved into Managed futures who were investors who invested in every market that was futures based such as equity indices, bonds futures, metals etc. Equity market investors also follow momentum investing and it can take two distinct forms owing to mandates given to equity portfolio managers.

There are different forms of momentum investing. One popular form is momentum analysis on cross-sectional data, such as components of FTSE 100. This means that the momentum is relative out-performance of one stock over another, or a set of stocks relative to rest of the constituents of an Index (Asness et al., 2013), (Hong and Stein, 1999) (Liu and Zhang, 2008). Our momentum work is different from cross-sectional momentum we are interested in absolute momentum in FTS data, since we are interested in identifying information in time series data (returns) at different time horizons.

Momentum investing works on the premise that securities that have performed poorly will continue to do poorly and the ones that have performed well will continue to do well over the long run. In essence such models are looking for returns with the same sign in financial returns data, that is to say they are looking for the development of a trend. Trends have been attributed to behavioural biases such as investor overconfidence, investors reacting or processing information slowly (De Bondt, 1985). Most importantly the existence of these trends has been shown to be statistically significant by Lemperiere et al. (2014) through an in-depth study looking at prices going as far back as 200 years. Several other researchers have also shown momentum to exist in several asset classes including commodities, equities, FX and equity indices (Moskowitz et al., 2011). However, some researchers have found that momentum investing may get some support from survivorship bias, especially in uptrends. Nearly all major indices such as FTSE 100 rebalance their indices on a regular basis removing the worst performer and adding the new stronger performer, hence creating a small bias towards uptrends (Henker and Huynh, 2010).

Momentum focused portfolio managers use many technical tools to identify trends but one of the most widely used method is the moving average and its many variants (Covel, 2007) (L'habitant, 2007). Some of the well known methods are a) simple moving average, b) moving average convergence divergence, c) exponential moving average, d) triple moving averages, e) adaptive moving averages, f) variable length moving averages, g) fixed length moving averages, h) high low moving averages. We will use some of these methods for our model as well in the next chapter.

To build a momentum model one needs to conduct some statistical tests that show trends exist and the returns can have the same sign from one period to another. One also needs to be able to extract a signal from noisy data. We show the tests that we used in Chapter 3.

2.3.2 Equity Market Neutral

We now turn to the market neutral case. Market Neutral (MN) means having little to no correlation with a broader market or the key index such as FTSE 100, which represents the broader UK market. The goal of a market neutral portfolio manager (PM) is to reduce or minimise market risk as much as possible while generating positive returns independent of the broader market. Selling or buying of securities is not a single stock decision, but related to other stocks in a well structured approach, such that some part of risk is reduced or eliminated.

In financial markets to profit from a fall in price of stocks, a PM needs to be *short* that stock. For example a PM can borrow Stock A from a bank, sell it in the stock exchange at value of say $10 with the expectation that the price will decrease. When the price of the stock decreases to say $9.50, the PM can buy the stock and return it to the bank and earn a return of 5% minus some costs. Similarly a PM going *long* is the term used to buy stocks with the objective of earning a profit when the price rises. A PM needs to be able to go both *long* and *short* to construct a MN portfolio. A MN model can take many forms and this can depend on the detail that PM is willing to go to, as well as the PM's investment style. Some of the well known ways for a PM to be market neutral are:

- **Dollar Neutral**: Dollar neutral stands for equal capital allocation. A PM who is long and short on equal amounts of capital on the underlying asset such as stocks, bonds or futures. Where *long* is equal to buying an asset and *short* is equal to selling an asset. For example a PM with $10 million in long positions would have to be matched with $10 million in short positions, with the PM expecting the long positions to increase in value and the short positions to decrease in value, giving the PM a profit. Dollar neutrality is extremely appealing because of its simplicity and ease of management for a PM on a day-to-day basis.

- **Beta Neutral**: A commonly used risk-based definition of market neutrality relies on Beta. A portfolio is said to be market neutral if it generates returns that are uncorrelated with the returns of some market index such as FTSE100. Since Beta is calculated by regressing stock returns to a market index, showing correlation, a zero correlation with the market implies zero Beta. To create a Beta Neutral portfolio we have to go back to MPT (see Section **??**). According to MPT the movement or volatility of a stock can be explained by the market risk (Market Index) and company (Stock) specific risk component (see Figure 2.3). The market risk component is driven by the market volatility. To construct a Beta Neutral portfolio the weighted average of stock Betas on the long side and the weighted average of the stock Betas on the short side must cancel each other out. For example a PM may be *long* on a set of stocks that have a weighted Beta of 0.85 and *short* a set of stocks that have weighted Beta of -0.85 with the Betas effectively cancelling each other out, inducing Beta neutrality. Depending on the variance of the stocks in the portfolio the PM may have to rebalance the portfolio from time to time. This usually depends on the mandate as well the the time horizon on which the PM is operating.

- **Sector Neutral**: Even though Beta Neutral portfolios are theoretically market neutral, sector specific movements can have an impact on the portfolio. For example a major change in insurance regulation can orchestrate a big move up or down in the sector compared to the rest of the portfolio. Sector neutrality can protect a portfolio from sudden major move owing to unanticipated changes in law or regulation which in most cases cannot be hedged. Sector neutrality would strengthen a Beta neutral portfolio, especially with control for market capitalisation, where market capitalisation stands for the market value of a company's equity. Hence a sector neutral PM would have the same amount of sector exposure on the long

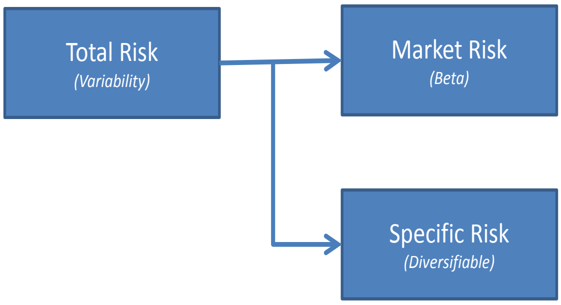


FIGURE 2.3: In this figure we show that their are two broad compenents to risk, the first is broad market risk usually attributed to the major national index such as FTSE 100. The other is stock or company risk which is specific to each comapny. Source: L'habitant (2007)

and short side of their portfolio. For example, a PM may be interested in a sector that has 10 stocks. Then the PM can be *long* 5 stocks that have a weighted Beta of 0.75 and *short* 4 stocks with weighted Beta of −0.75 and may exclude 1 stock. This is achieved through slightly different capital allocation to the long and short side to achieve Beta neutrality since Beta is calculated as a weighted average. This is how a PM is Beta neutral and sector neutral.

- **Factor Neutral**: Factor Neutral portfolios are some of the more complex portfolios to build and require extensive fundamental information on stocks. By factors we mean micro factors such as earnings, price-to-book ratio, dividend yield, which are company specific as well as macro factors such as oil price, interest rates and inflation. BARRA, a long established commercial entity specialises in providing analytics, identifies as many as 68 micro and macro factors for the US stock market. Here we decompose the market variability further into factors i.e. we get more detailed than just Beta neutrality shown in Figure 2.3 to Figure 2.4.

  It is important for the PM to identify factors that impact returns. One way of finding useful factors is to rank the stocks by a certain factor, such as book value, which means the true value of a company if all its assets were sold today. Taking the difference between the average return of the top quartile and the bottom quartile, if the return is positive on say a monthly or weekly frequency then we have identified a useful factor that can be used. Using this methodology we can then find other factors and combine them and build a multi-factor model.

  A PM requires a sophisticated and detailed model to find relevant and useful common factors, identify precise source of risk in their portfolio, quantify them, so the PM can neutralise risk factors while have some exposure to desirable factors. Theoretically a PM can hedge all the factors (all 68 factors from BARRA) by being *long* and *short* same factor exposure and essentially earn the risk free rate of return after costs. The challenge for a PM is to identify the undesirable factor risk and desirable factor risk. Once a PM has identified the desired factors exposure, then the PM can build a market neutral portfolio gaining exposure to desired factors while hedging factors it may dislike. For example, most PMs deem it prudent to invest in companies that are cheap or undervalued according some metric of their choice, while minimise exposure to the oil price. A PM may decide to go *long* on stocks that score high and are cheap on the basis of book value, earnings and dividend yield, but have negative exposure to oil and short ones that score low and are expensive and have positive exposure to oil and expect to make a positive return, while still maintaining Beta neutrality (L'habitant, 2007).

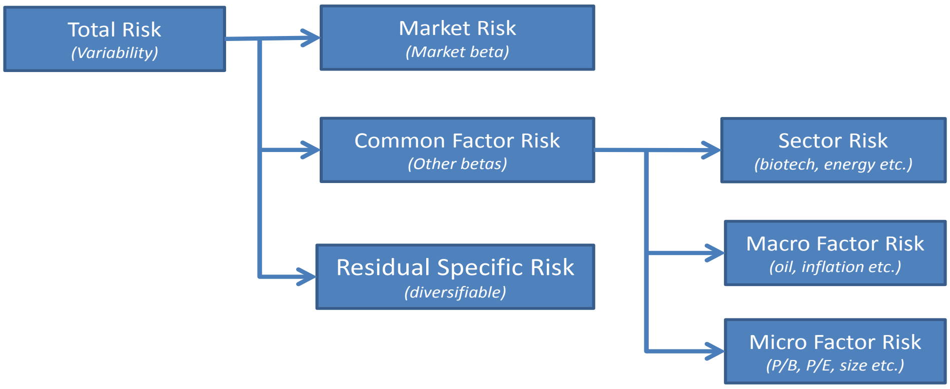


FIGURE 2.4: Factor Neutral models need to look at factor risk is much more detailed manner to identfy every little driver of variability, as shown in this figure. Source: L'habitant (2007)

- **Pairs Trading:** Pairs trading, as the name implies, involves two securities. They can be two bonds, stocks, index futures or commodities. More specifically, pairs trading involves two related securities with similar characteristics that move together and are likely to deviate temporarily from their long term path. Therefore, when they deviate far enough from a historical or statistical perspective they generate an expectation that they will revert back to their historical levels. The approach is based on mean reversion, which is essentially making a call on the relationship between two stocks. The process of pairs trading can be simplified to a few key steps:

  - Identify pairs of stocks where prices should move in tandem;
  - take a long and a short position in these related stocks when their prices diverge sufficiently;
  - hold the position until the prices have converged back to their normal relationship range, or until they hit a pre-set stop loss level.

  There are many methods to perform pairs trading. One method is to take each security as standalone and forecast them using a model. Another method is to carry out fundamental analysis of two securities, alternatively there is the approach of creating a *spread* between the price time series of the securities. Here, the *spread* is the log difference of the price time series of the relevant securities. The *spread* represents the price of a security relative to another security.

Pairs trading in equities (stocks) is nearly always done in a sector neutral manner to get relatively stable spreads, so they can be modelled well. Stocks especially of large corporations have very standardised businesses and regulatory environment, hence have similar factors exposure. For example, a PM may want to build a portfolio of pairs in the UK. The first step is to take all the stocks and categorise them into sectors. The next step is to select stocks that represent large, well established companies such as those in the insurance sector. Finally one needs to check if their business revenue streams come from similar sources e.g. is it life insurance, car insurance or general insurance.

Most models use some kind of a distance measure with a threshold. For example, this threshold can be two standard deviations away from the mean as estimated using lagged observations or a certain percentile measure of the empirical distribution. A position is opened when the distance threshold is breached and closed when another threshold is reached; either with gain i.e. mean reversion has occurred or with a stop-loss when a position has not converged. There are other distance measures such as co-integration, stochastic spread approach with Ornstien-Uhlenbeck model or through orthogonal regression as discussed by Elliott et al. (2005), Vidyamurthy (2004) and Gatev et al. (2006). Gatev performed a study of stock pairs in the US market including sector neutral pairs, using data from 1962 to 2002. He used the the top performing pairs, which showed promising results and built pairs portfolio showing good results. We will use the stochastic spread approach using Ornstien-Uhlenbeck model. We will present tests that we conduct in Chapter 3.

- **Statistical Arbitrage:** Statistical arbitrage is an extension of pairs trading, with factor exposure. In statistical arbitrage analysts and PMs consider baskets of stocks, rather than pairs (however people use the term loosely for pairs as well). In more detail, they divide securities into different groups based on several criteria and look for systematic divergences between these groups. Their portfolio will typically consist of a large number of long and short positions chosen simultaneously; for instance, they may buy the 20 percent most undervalued (cheap) stocks and sell short 20 percent most overvalued (expensive) according to some criteria such as book value, with the aim of capturing the average mis-pricing between groups corrects.

  Statistical arbitrage can be seen as an extension of the pairs trading approach to relative pricing. The underlying premise in relative pricing is that groups of stocks having similar characteristics should be priced on average in the same way. However, due to non-rational, historical or behavioural factors, some discrepancies

> may be temporarily observed. Rather than looking for a few pairs of securities that diverge from their historical relationship, statistical arbitrageurs slice and dice the whole universe of stocks according to sectors, valuations, factors or a combination of these categories.

With regards to the market neutral approach for our research and model development we have focused on pairs trading with dollar and sector neutrality, as we have access to relevant data. We shall discuss the tests we did in Chapter 3 and performance of the model on real data in Chapter 8.

2.3.3 Long Only

Long only portfolio management or buying and holding several securities to build a portfolio has been the approach most equity portfolio managers have used to manage money. Even today the largest investments, in terms of the amount of money invested are in long only investments around the world. This stands to reason as it is easier to buy securities for the long run and hold them. A lot of portfolios are index trackers, that is to say they essentially track the major indices such as FTSE100, S&P 500 etc. PMs take decisions to buy and sell stocks either by reading research reports, through quantitative investment models or a mix of both, holding stocks for varying amounts of time. PMs using quantitative models tend to use multi-factor models that utilise accounting information such as price-to-sales ratio and book value, as well as pure time series models to identify momentum in prices, where both approaches are trying to identify stocks to buy. The objective is the same: to build a diversified portfolio, broadly representing most major sectors such as pharmaceuticals, insurance, telecom, mining banks etc. of a country's economy.

We will use a time series based approach to buy stocks. Specifically we will build a model that uses price time series, which is segmented by sectors, as stocks in the same sectors exhibit similar behaviour not limited to variance. Just as a diversified portfolio we will invest in all the major sectors represented by our model. We will show the tests that we use for this model in Chapter 3.

2.4 Meta Model of Quantitative Investment Models

In the previous section we discussed model based approaches that capture different aspects of price data on different time scales, namely momentum drive investing, long only and market neutral. As discussed in Chapter 1, our objective of building these models is to combine them into a portfolio; a portfolio of such models has many benefits. Firstly, we are able to capture more variance in the data, hence generate better returns. Secondly, since we capture more patterns in the data on different time scales, we can diversify our investment risk across different time horizons. Using the data in a encompassing manner gives a sense completeness to a portfolio manager's mandate. Combining these models is our key task and the next step we focus on.

In the hedge fund sector there are a number of strategies operating on different asset classes and generating returns is not entirely dependent on positive returns from the markets. Given the niche that we are looking at, there is limited literature specifically focusing on this task. For our objective the most relevant academic research was done by Burgess (1999) and there has been some empirical research also done by Amenc and Martellini (2002). Burgess in his PhD thesis selects few optimal statistical arbitrage (SA) models from several similar models, while Amenc and Martellini (2002) uses the longest available database of hedge fund performance from Tremont/Credit Sussie. Burgess selects the best models from a population of 270 statistical arbitrage models, while Amenc and Martellini use nine hedge fund strategies to form a portfolio.

To build their portfolio Amenc and Martellini (2002) use the mean variance optimisation framework, which we think is inappropriate given the data generated by hedge fund strategies. The model devised by Burgess is very different. He uses a population based model. The objective of this is to maximise the risk-adjusted return as measured by the Sharpe ratio while controlling for correlation. The population models works in the following manner:

- Generate candidate SA models.
- For each SA model generate a meta parameter in the form of Sharpe ratio.
- Identify SA models that have the highest Sharpe ratio to start with.
- At each time step add models that either increase return, reduce risk, or reduce correlation in the meta portfolio
- Stay with previous meta portfolio if there is no value addition.

Specifically Burgess is building an optimal portfolio, while controlling for correlation and optimising to maximise the Meta portfolio's Sharpe ratio, without using MVO. However he should get the same result as MVO asymptotically for correlated assets. In MVO during optimisation, the covariance matrix captures the joint distribution of returns instead of explicitly doing this Burgess controls it through correlation.

Burgess's work is very encouraging and has similarities to our work as well as some differences. All of Burgess's models are quantitative just like ours and similar to our approach he wants to combine them, as we show in Chapter 8. However, there are some key differences. Burgess does not build models to capture data structure at different time horizons and only focusses on SA models, which can have strong positive correlation. Specifically Burgess focuses on 270 optimised SA models and chooses the best from his sample set. Since most SA models are built in a similar fashion and have very similar risk return profile, a population based model would make sense. However, we build models that are more diverse, and we categorise them into broader groupings such as momentum and market neutral (Section 2.3), while Burgess has one approach, that of SA.

Burgess's approach would not be appropriate for models capturing structure and patterns found at differing time horizons. Furthermore Burgess pays no attention to investment size, simply focusing on Sharpe ratio. As we discussed in Section 2.2.4 it is not prudent to over bet even when you have positive expectations.

Balvers (2006), who had been working on momentum and mean reversion models previously, decided to build a model that combined the two components. Balvers *et al* build on the research of Jegadeesh and Titman (2001) and Lee and Swaminathan (2000). A key observation made by Jegadeesh and Titman (1993) that portfolios built using a momentum model eventually experienced mean reversion a few months later. They built a single model, by decomposing the returns of momentum and mean reversion, using the methodology of Fama and French (1988) and Summers (1986) and then combined them into a single model.

Balvers *et al* used a very large data set of 18 international markets from 1977 to 1999, at monthly frequency. They report better risk-adjusted returns than his previous models that focused on momentum. Serban (2010) applied Balvers *et al* model to the FX market and also reported higher risk-adjusted returns, when compared to either a mean reversion or momentum model for the FX market.

Balvers *et al* work is similar to ours in spirit. That they want to capture the two most prevalent patterns in financial data, mean reversion and momentum at differing time horizons. However our approach is different, they have built one model that incorporates both momentum and mean reversion to build a single portfolio, whereas we want to build dedicated models for each pattern and then make a portfolio of models. We also want to address interactions between related data which they don't, nor do they address the important point of investment size. Furthermore their data set is also focused on equities but they have much longer history than ours but only on monthly frequency. Even though our research diverges at a certain point from Balvers and Serban, the results that they report are very encouraging for us and we aim to achieve better results with our model as well.

2.5 Kalman Filter

In the previous section we discussed a number of QIMs. All of them rely on tools from statistics and control engineering to make them robust. One of the tools used in finance is the Kalman Filter. Financial data is hard to model well, as the data is noisy and non-stationary, as we discussed earlier in Section 2.1. To make modelling such data manageable analysts have borrowed tools from different fields, such as statistics, control engineering signal processing and Physics to name a few.

The Kalman Filter is a recursive predict-update algorithm devised by Kalman (1960), (Kim, 2011). This algorithm comes from control engineering and is used extensively in machine learning as well. The Kalman Filter has been used in finance for various purposes, such as volatility models, to estimate option price, overcome outliers in data as well as reduce estimation error owing to noisy data, smoothing of times series data for stocks, estimating missing values on volatility surface extrapolated from option prices etc. smoothing estimation of stock Beta calculations (Javaheri, 2005), (Wells, 1996). There are variants to the Kalman Filter such as Unscented Kalman Filter and Extended Kalman Filter used in finance and Javaheri discusses them in some detail with reference to applications in finance.

We briefly recall the steps of the Kalman Filter for linear state-space models. We introduce the following notations:

$$\theta_t = \mathbf{A}_t \theta_{t-1} + w_t \qquad w_t \sim N(0, W_t) \qquad . \tag{2.10}$$

$$z_t = \mathbf{H}_t \theta_t + v_t, \qquad v_t \sim N(0, V_t). \tag{2.11}$$

$z_t \in \mathbb{R}^m$ stands for the observation vector, $\theta_t \in \mathbb{R}^p$ is a hidden random vector, $\mathbf{H}_t$ is the observation matrix; and $\mathbf{A}_t$ is the system matrix, that predicts our position at the next time step; and are of size respectively $(m \times p)$ and $(p \times p)$, to be specified, and V_i and W_t are the observation and evolution covariance matrices of size $(m \times m)$ and $(p \times p)$ respectively.

In the following sections, we assume that $V_t = V$ and $W_t = W$, for any t. They are estimated from available in sample data. The Kalman Filter recursively estimates the internal state of the process θ_t given the sequence of noisy observations z_t. We denote by θ_t the estimate of the state at time t given observations up to and including time T, and by P_t the associated error covariance matrix. This can be summed up by the system of equations:

$$\hat{\theta}_{t|t-1} = \mathbf{A}_t \hat{\theta}_{t-1|t-1}. \tag{2.12}$$

$$\mathbf{P}_{t|t-1} = \mathbf{A}_t \mathbf{P}_{t-1|t-1} \mathbf{A}_t' + W_{k-1}. \tag{2.13}$$

$$y_t = z_t - \mathbf{H}\hat{\theta}_{t|t-1}. \tag{2.14}$$

$$\mathbf{S}_t = \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t' + V_t. \tag{2.15}$$

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t' \mathbf{S}_t^{-1}. \tag{2.16}$$

$$\hat{\theta}_{t|t} = \hat{\theta}_{t|t-1} + \mathbf{K}_t r_t. \tag{2.17}$$

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1}. \tag{2.18}$$

Equation 2.12 gives the predicted state at step t and Equation 2.14 the innovation residual, where the innovation is the difference between the observed value z_t of a variable at time t and the optimal forecast of that value θ_t based on information available prior to time t. S_t in Equation 2.15 is the innovation covariance and compares the real error against prediction. K_t in Equation 2.16 is the the Kalman gain which moderates the prediction based on the accuracy of the last time step $t-1$ and Equation 2.18 represents the new estimation of error for the next time step (Mahler, 2009). We will use the Kalman Filter in our models and discuss its application in Chapter 5.

2.6 Summary

In this chapter we presented an overview of the previous research and literature that has studied the characteristics of financial data and we saw the challenges this poses for analysis and development of quantitative investment models. We then looked at various methods of constructing portfolios, their benefits and shortcomings as well as applicability for our research. We highlighted the power of the mean-variance framework and its intuitive appeal, the evolution of MPT with CAPM and APT and their variations such as the Black-Litterman model. Subsequently we discussed some newer models such as portfolio from sorts, universal portfolio, and risk parity approach. We then discussed Kelly's model for betting and how it can be used to build a portfolio. We also show several very desirable properties of Kelly's approach.

The quantitative investment model section presented the widely used and well established models in the equity markets. We discussed their many variations and what they achieve. We will use some of them to capture momentum and mean reversion in data. We then discussed some previous work done on making portfolios out of quantitative investment models particularly that of Burgess. We also discussed previous work done on capturing momentum and mean reversion, especially by Balvers. We also presented the Kalman Filter which we use in some of the models.

This chapter sets a sharper and clearer background for our research. Some on the models presented in this chapter will used as benchmarks by us to evaluate our results later. Some models from the quantitative investment model section will be implemented to construct investment models that capture momentum and mean reversion with high probability of success, which as discussed in Chapter 1 is one the important ingredients of good portfolio management. In the next chapter we will present some of the steps and tests that we use to test for momentum and mean reversion in our data.

Chapter 3

Statistical Tests to Identify Structure and Patterns in Data

In the previous chapter we discussed existing work that is relevant to our research. We also discussed several quantitative investment models, especially ones that we will use for our research. To understand the nature of our data and to identify patterns and structure in the data, we have to do some statistical tests, which give us a path to building these models. In this chapter we will present statistical tests, which help identify patterns or structure in data, as well as how we can use some of the results from these tests to identify parameters for our models. Specifically, we will focus on tests relevant to the momentum model, market neutral model, as done through dollar and sector neutral structure and the long only model.

3.1 Statistical Tests and Analysis

In this section we discuss some of the tests that we performed to identify which indices and stocks, and which relationships among these stocks and indices, should be pursued to build an investment model. We are looking to conduct tests that will help us identify structure in data and establish relationships that should exist as described in financial and economic theory, as well as understand the general characteristics of our dataset. Some of the tests are generic statistical tests and some are useful to identify certain features that would help build a quantitative investment strategy. We will categorise them as such for clarity. The histograms we present are made using the the Freedman - Diaconis method to calculate bin size. We run our tests using MATLAB© version 2013A.

3.1.1 Empirical Distributions

We run some initial tests on the data to identify whether data is normally distributed and to identify some other statistical features in the data such as checking moments of the data namely mean, variance, skew and kurtosis. Normally distributed data does not have third or fourth moments called skew and kurtosis; i.e. they are 0.

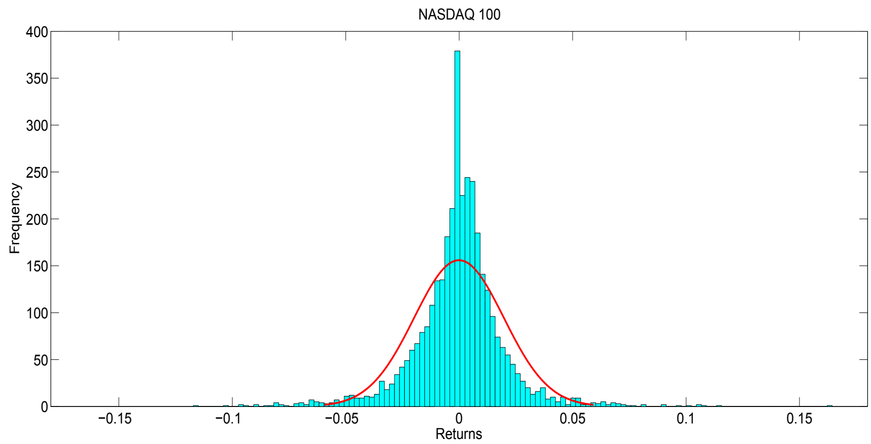


FIGURE 3.1: In blue we can see the empirical distribution of NASDAQ 100 futures contract at daily frequency, compared to the theoretical normal distribution curve in red, showing the data is not normally distruibuted and exhibits skew and kurtosis. Mean = −0.00007, variance = 0.01975, skew = 0.01440 and kurtosis = 9.01269. The bin size is calculated using the Freedman - Diaconis rule. Thebin size is 0.002217 and there are 126 bins.

Kolmogorov-Smirnov test for normality

The one sample Kolmogorov-Smirnov test compares the value of single data array z to a standard normal distribution, i.e. a normal distribution with zero (0) mean and unit (1) variance. The null hypothesis for the Kolmogorov-Smirnov is that z has a standard normal distribution. The alternative hypothesis is that z does not have that distribution and we reject the hypothesis if the test is significant at the 5% level. At each potential value of z the Kolmogorov-Smirnov test compares the proportion of values less than z with the expected number predicted by the standard normal distribution as shown in Figure 3.2.

We now check the distribution of our data, and the four moments, namely mean, variance, skew and kurtosis. In Figure 3.1 we present a specimen from real financial data. The distribution of the data shows that the data is not only heavy tailed, but it is also quite noisy. We can see that from the variance in the data.

In Figure 3.3 we can see how price time series data can exhibit drift or momentum as well as short-run mean reversion to moving mean depicted by a 10-period simple

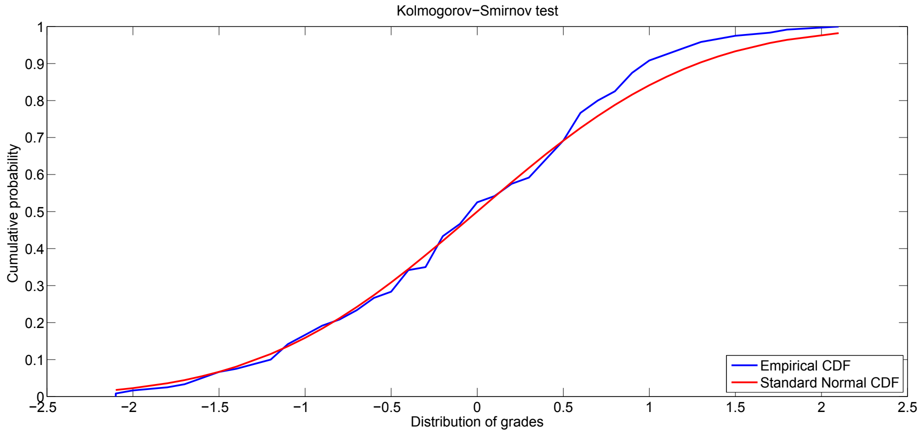


FIGURE 3.2: A specimen of Kolmogorov-Smirnov test for normality of data. This specimen shows that the data shown by the blue curve is not normally distributed, when compared with the red curve which is from normally distributed data. Hence the data fails the normality test.

moving average. The Software AG stock shows a near 50% appreciation in price over approximately 100 days. However it shows short-run mean reversion to moving mean as measured by a 10-period simple moving average.

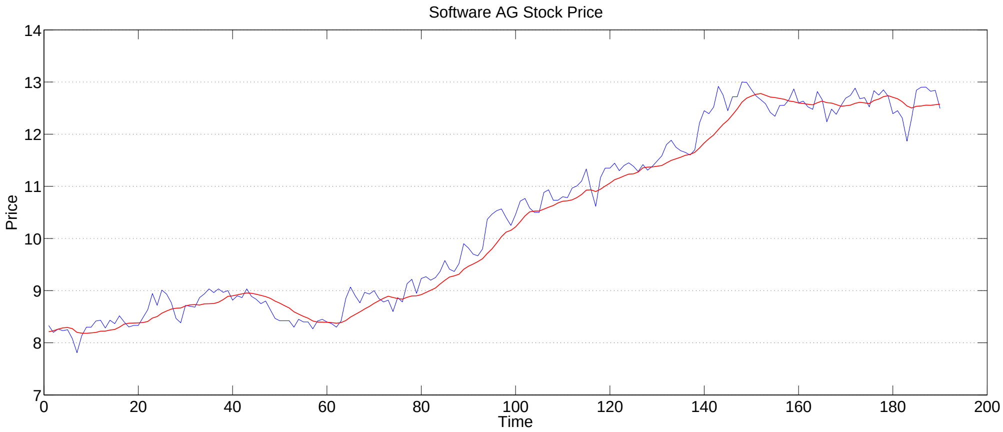


FIGURE 3.3: Software AG price from 14/1/2005 to 6/10/2005 exhibiting trend in the long run. Moving mean is calculated using a 10-period simple moving average, accompanied with noise that is mean-reverting to the moving mean of the stock price.

Our objective is to identify structure in financial data so that we can parametrise and build a QIM for that particular pattern or structure. To do this we need to organise our data in a practical manner. This is particularly relevant for the market neutral model that we intend to build using the pairs methodology, we outlined in Section 2.3.2 and the long only model discussed in Section 2.3.3. For both these models we will have to categorise the data so that it is relevant to the approach taken by the QIM.

3.1.2 Tests for Market Neutral Model

For the market neutral model we use the pairs methodology with dollar and sector neutrality (see Section 2.3.2). Since stocks in the same sector are highly correlated and have similar variance, we expect to see mean reversion among pairs from the same sector, that might have had some short-term dislocation, especially in the spread of the two stocks. We now present statistical tests done for this model and subsequently discuss how these steps help us build the market neutral model.

- **Regression analysis** We perform a regression analysis using ordinary least square (OLS) using Equation 3.1, on the two price times series that comprise all the pairs, and store the regression residuals ϵ_t as shown in Equation 3.1,

$$y_t = \beta x_t + \epsilon_t. \tag{3.1}$$

  By regression residuals we mean the unexplained part of the regression. These residuals are of interest to us as they show the nature of their relationship between two time series. We ideally want the residuals to be strongly mean reverting, crossing the mean very often. Pairs that have this behaviour are of most interest to us. We also check the coefficient (β) and expect it to be less than one or else it would signal an explosive process

  We then conduct the Runs test on the residuals to check the degree of serial correlation in the residuals, and expected mean reversion in the residue. Pairs of stocks that have their regression residuals cross the mean many times are of most interest to us, as they are likely to be the most stable and consistent mean-reverting relationships.

  This test is important for the pairs model as it helps us identify relationships we should pursue further. This test is also useful for sector analysis as, according to financial theory, stocks from the same sector usually have very similar variance. In Figure 3.4 on the top half of the chart, we show two stocks rebased to 100 with some trend taking them on slightly different paths. In the bottom half we show the difference between the two stocks, which is called the *spread*, such as $\log(StockA) - \log(StockB)$

- **Runs Test** The Runs test is designed to detect serial correlation in univariate time series (Bradley, 1968). We use this test on the residuals (error term) from the regression we did above, where the residuals are the unexplained portion of a regression. Through the Runs test we count sequence of positive and negative values above and below a mean of the residuals, and we run the test on independent windows (no overlap) of 60 and 120 days as well as the whole in-sample dataset. We want the residuals to cross the mean as many times as possible (Figure 3.5),

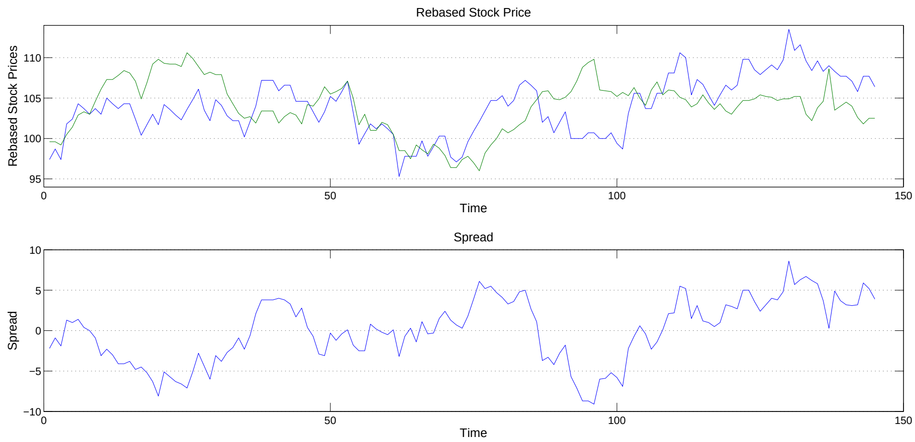


FIGURE 3.4: A specimen of two re-based stocks. Hannover Reinsurance (Germany) and SCOR (France) from 01/02/2005 to 22/08/2005 are shown on the top plot. The spread is the difference between the two prices shown in the bottom plot. Such a spread is the ideal model for a pairs trading approach.

which would indicate that the relationship is mean-reverting, hence identifying a suitable relationship to use in our model.

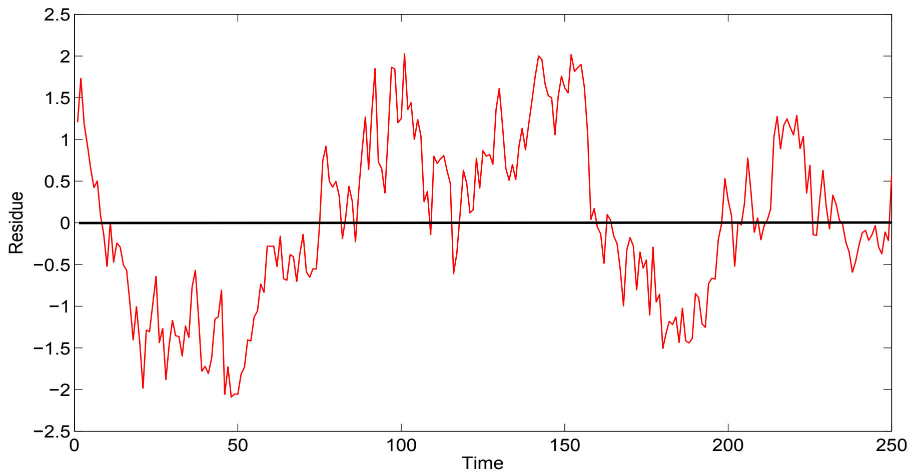


FIGURE 3.5: Example of mean-reverting regression residue, which crosses the mean consistently. Such behaviour is a positive sign for a mean-reverting relationship.

- **The Variance Ratio Test** (VRT) is a non-parametric test for randomness in time series data and we use it to test to see whether a price time series is a random walk, since in some studies it has been found to yield good results (Lo and MacKinlay, 1988). Specifically the VRT tests whether variance in shorter windows matches to longer windows, with the null hypothesis being that the variance ratio is 1, since variance should linearly scale from shorter time frame to longer time frame. VRT values below 1 indicate mean-reverting behaviour and at 0.5 it represents perfectly random data. This is important as we will be modelling mean-reverting process using the Ornstein-Uhlenbeck model.

In Equation 3.2 we outline the formula for the test,

$$VR(\tau) = \frac{\sum_t (\Delta^\tau y_t - \overline{\Delta^\tau y})^2}{\tau \sum_t (\Delta y_t - \overline{\Delta y})^2}. \tag{3.2}$$

Here τ is the length of the long-term variance, y_t is the time series in levels, Δy_t is the daily change in the time series $(y_{t+1} - y_t)$, $\Delta^\tau y_t$ is the long-term change in the time series $(y_{t-\tau} - y_t)$, $\overline{\Delta^\tau y}$ is the mean value of the long-term change and $\overline{\Delta y}$ is the mean value of the short term change

To understand the variance of data we create a a Variance Ratio Profile (VRP). The VRP is the result of putting together several VRT on different time horizons, as seen in Figure 3.6. For example we create VRP over a period of 20 days starting from 2 days to 20 with 20 days approximating one month. This helps us to identify whether a data series on this time horizon is mean reverting or trending. We will use this for one of our market neutral models.

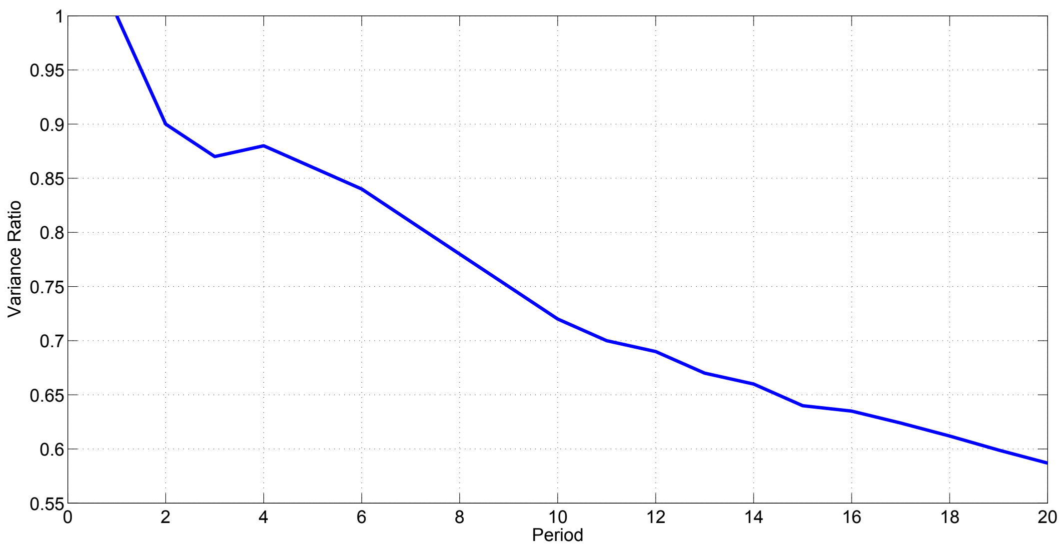


FIGURE 3.6: Specimen of a variance ratio profile made from variance ratio 20 periods. A falling variance ratio profile shows that data is mean reverting in nature.

Having done the above mentioned statistical tests, we find spreads from paired data that have potential to be part of the market neutral model. The spread that shows potential is modelled using the Ornstien-Uhlenbeck (OU) model shown in Equation 3.3

$$dS_t = \lambda \left(\mu - S_t\right) dt + \sigma dw_t. \tag{3.3}$$

Here S_t is the *spread*, t is time, μ is the mean of the spread, λ is the speed of mean reversion, σ is volatility, w_t is a stochastic term. How we build the model as well how the model operates, we will discuss in Section 5.1. In discrete case we ca an ordinary least square setting such as Equation 3.4,

$$y = \alpha + \beta x + \epsilon. \tag{3.4}$$

Solving Equation 3.3 for time period δ we find that $S_{t+\delta} = e^{-\lambda\delta} S_t + \mu(1 - e^{-\lambda\delta}) + \epsilon$. Where $E(t) = 0$ and $E(\epsilon^2) = \sigma\sqrt{\frac{1-e^{-2\lambda\delta}}{2\lambda}}$. Fitting a regression line to a time series S_{ti} measured at time interval δ with a fit $S_{t+i} = aS_t + b + \epsilon$ we find $a = e^{-\lambda\delta}$ or $\lambda = -\frac{\log a}{\delta}$. λ becomes a distance measure, the further the current observations is from the average (μ) the larger λ gets an the higher is the speed of mean reversion.

The null hypothesis is set in reference to the VRP, where the null hypothesis is that there is no mean reversion. The null hypothesis is rejected when the VRP score is at or below 0.5, showing mean reversion. We will use these tests in Chapter 5, where we build quantitative investment models.

3.1.3 Tests for Momentum Model

In the momentum model we want to identify time series that show continued movement either up or down for long periods. To build this model we first need to identify if the data has returns that have the same sign i.e. trends. The initial step is to change the frequency of the data from daily to monthly. We make this transformation to the data for two reasons. Firstly to reduce the noise in the data to get reliable estimate, and secondly, to identify long-run trends in the data.

- **Autocorrelation function** This function is the internal correlation of the observation in a time series usually expressed as a function of the time lag between observations. This autocorrelation function generally gives better results when noise in the data is low. It is normally best to use the autocorrelation function on lower frequency data. Hence we do the test for autocorrelation on monthly frequency data. In Figure 3.7, we show the autocorrelation plot for monthly returns of NASDAQ 100 futures contract, which shows dependence only to one lag.
- **Auto Regressive Integrated Moving Average Model (ARIMA)** The ARIMA model is a well known approach for modelling stationary as well as non-stationary time series data introduced by Box and Jenkins in 1968 (Box and Jenkins, 2008) and is based on the the acceptance that any stationary time series can be approximated by a combination of autoregressive (AR) and moving average (MA) processes, so called ARMA processes i.e. of the form:

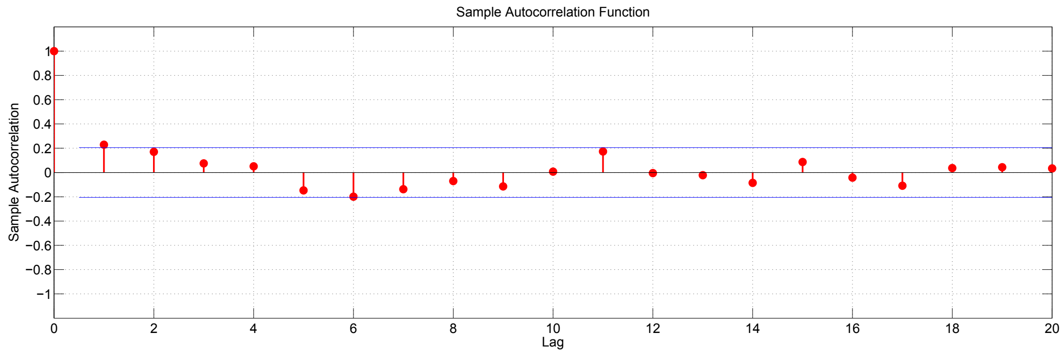


FIGURE 3.7: Autocorrelation function plot of monthly returns of NASDAQ 100 futures contract. The data in the in-sample period (2005 – 2009), shows autocorrelation in returns data, significant only to one lag.

$$y_t = \mu + \sum_{i=1..p} \phi_i y_{t-i} + \sum_{j=1..q} \theta_j \epsilon_{t-j} + \epsilon_t. \quad (3.5)$$

Here in Equation 3.5 $\phi_i y_{t-i}$ is the AR term, ϕ_i is the coefficient of the AR term, $\theta_j \epsilon_{t-j}$ is the MA term, θ_j is the coefficient of the MA term, μ is the mean and ϵ_t is the residual term. A model with $p>0, q = 0$ is denoted as AR(p), and is dependent only on its lagged values and is referred as a pure autoregressive model. A model with $q>0, p = 0$ is denoted as MA(q) and is dependent on lagged values of innovations, referred to as a pure moving average model. For non-stationary time series, data needs to be differenced to make it stationary; the number of times the data needs to be differenced to make it stationary is referred to as its *order of integration*. Data that is differenced once to make it stationary is said to be integrated to the order 1. In ARIMA, the I stands for the order of integration, an ARIMA(p, 1, q) models is one that needs to be differenced once (see Equation 3.6),

$$y_t = \mu + \sum_{i=1..p} \phi_i \Delta y_{t-i} + \sum_{j=1..q} \theta_j \epsilon_{t-j} + \epsilon_t. \quad (3.6)$$

- **Regression Analysis**

  To identify momentum or trend we perform regression analysis. We follow Moskowitz et al. (2011) approach where they performed a study of momentum across asset classes spanning a long horizon. They performed regression analysis to identify momentum or trend using monthly price time series data.

  We perform regression analysis to test for predictability of future returns based on lagged returns. We regress returns r_t^s for equity index s in month t on its returns lagged h months. Where both returns are scaled by their ex-ante volatility σ_{t-1}^s, as shown in Equation 3.7

$$r_t^s / \sigma_{t-1}^s = \alpha + \beta_h r_{t-h}^s / \sigma_{t-h-1}^s + \epsilon_t^s. \quad (3.7)$$

We run regression up to 12 lags. We get positive coefficients and corresponding positive t-statistic for up to 5 months as shown in Figure 3.8 and Figure 3.9. According to Moskowitz et al. (2011) positive t-statistic shows significant return continuation or trend. Negative signs indicate trend reversal. This analysis was performed by Moskowitz et al. on a large scale study of time series momentum across commodities, bonds, equities and foreign exchange. The size of our data set is considerably smaller than the one used by Moskowitz et al. This is probably why we observe two of the datapoints to be above the critical threshold at 0.05 significance level as indicated by the red line in Figure 3.9.

Another approach used by Moskowitz et al. to look at time series predictability is to simply focus on the *sign* of the past returns as shown in Equation 3.8,

$$r_t^s/\sigma_{t-1}^s = \alpha + \beta_h Sign(r_{t-h}^s) + \epsilon_t^s. \tag{3.8}$$

Just as in Equation 3.7 the returns are scaled by their ex-ante volatility making the left side of the regression independent of volatility and the right side is too, since *sign* is either -1 or $+1$. We can see a specimen of the same universe of 17 markets in Figure 3.10 and Figure 3.11. In terms of coefficients the results are similar to the previous regression as can be seen in Figure 3.10. However for this regression the t-statistics are below the critical threshold at 0.05 significance level as seen in Figure 3.11.

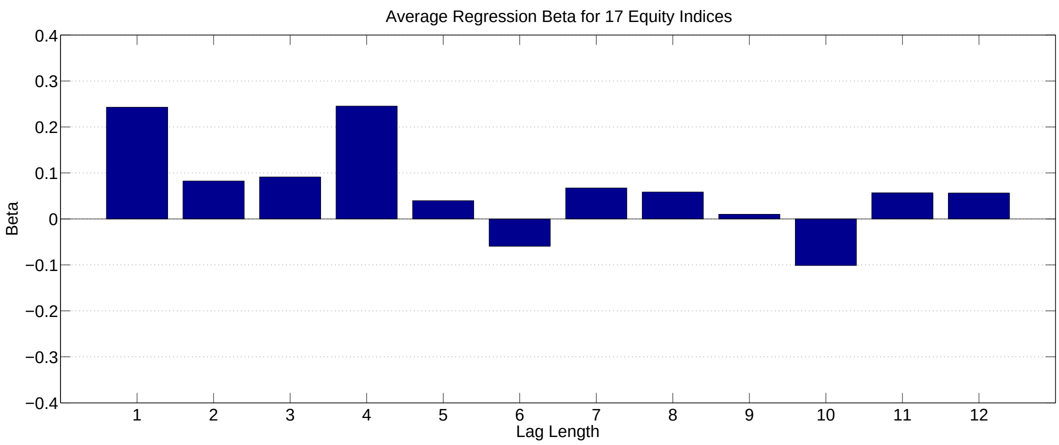


FIGURE 3.8: This figure shows the average Beta of 17 equity indices. Positive Beta with positive t-statistic signals that there is predictability up to five-month lag.

The null hypothesis is set with reference to the regression analysis. The null hypothesis is that there is no predictability in returns from lagged returns, while a positive regression coefficient with corresponding t-statistics would reject the null hypothesis. We will use these tests in Chapter 5, where we build quantitative investment models.

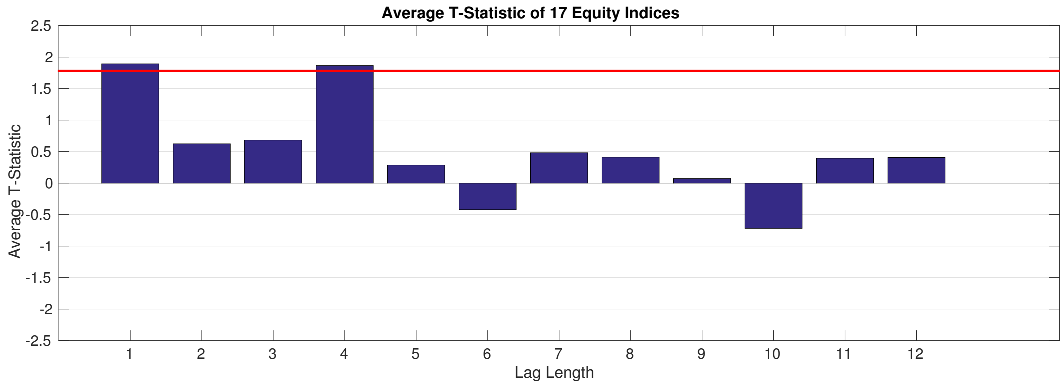


FIGURE 3.9: This figure shows the associated average t-statistic of the Beta for 17 equity indices. The red line represents the critical value threshold at 0.05 significance level. We can see that two values marginally breach the threshold.

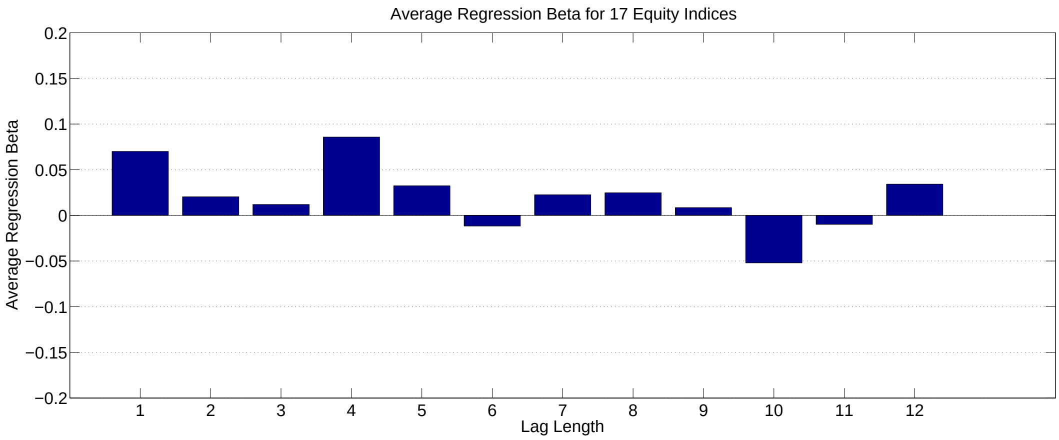


FIGURE 3.10: This figure shows the average Beta for the sign regression of 17 equity indices. Positive Beta with positive t-statistic signals that there is predictability up to five-month lag.

3.1.4 Tests for Long Only Model

The long only model is designed to only buy stocks (Section 2.3). In many ways large well established business are standardised as in the case of electricity or toothpaste; similarly the stock price movement reflects this standardised behaviour. Hence stocks from the same sector and of same market value should move together in terms of variance. For the long only model, just as in the case of pairs approach for the market neutral model, we have to create some grouping that resembles a sector in real financial markets. For example for the FTSE 100 Index the insurance sector has companies such as Royal Sun Alliance (RSA), Aviva (AV/), Standard Life (SL/), Prudential (PRU). As shown in Figure 3.12, all these stocks reflect very similar volatilities.

For this model we create custom sector grouping; e.g. insurance group will consist of equally weighted insurance stocks. We do this for each sector creating a custom sector

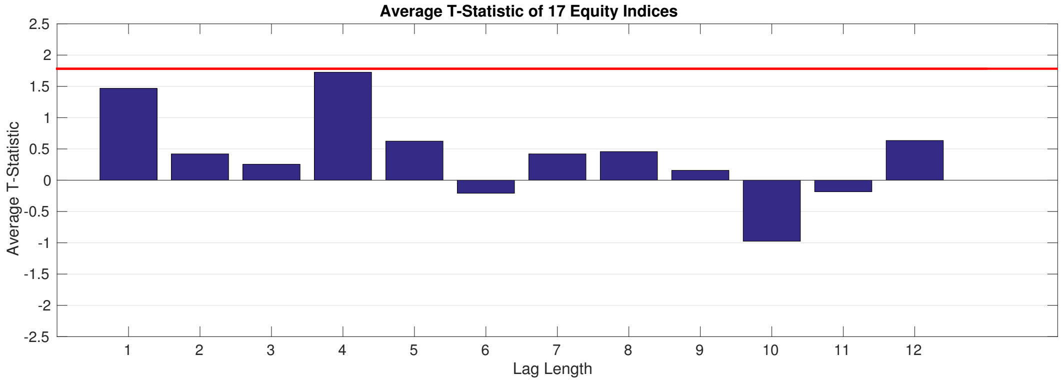


FIGURE 3.11: This figure shows the associated average t-statistic of the Beta for the sign regression for 17 equity indices. The red line represents the critical value threshold at 0.05 significance level.

index, giving us the sector average. Subsequently, we regress each stock on its respective sector index. We expect to see that each stock has a significant relationship with its sector, which we found in most of the stocks. We then do the Kruskal-Wallis one-way analysis of variance by ranks. This is a non-parametric test that compares variance to see if it comes from the same distribution. We want to see if our data comes from the same distribution. The test conducted by converting the data to ranks is shown in Equation 3.9,

$$\left(\frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{n_j}\right) - 3(N+1). \tag{3.9}$$

Here N represents all the observations in the data set and R^2 is the squared sum of the ranks. The test statistic has χ^2 distribution and degrees of freedom are the number of data vectors minus 1. This also identifies the significance level we want for the test (Hollander, 2014). The Kruskal-Wallis supports the economic theory of approximately similar distributions and hence variance.

We then follow this up with an analysis of the volatility of each stock in their respective sectors, (see Figure 3.12). As expected, we find that stocks in the same sector have very similar volatilities even during a crisis, in line with financial theory. In Table 3.1 we can see that 10-period volatility of the Insurance sector has strong correlation as expected. This is important as we expect stocks from the sector to move in tandem, making our model stable.

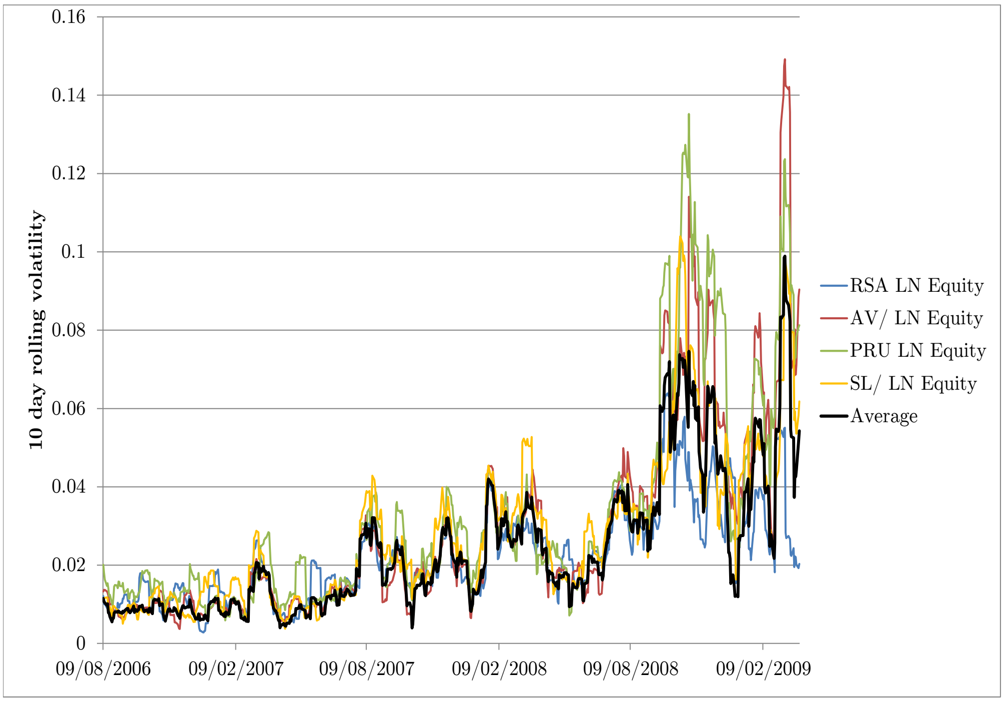


FIGURE 3.12: Specimen of four companies in the insurance sector of FTSE 100 Index. This figure shows Royal Sun Allicance, Aviva, Prudential and Standard Life and their average in black. Volatility of stocks in the same sector is very similar and hence their price movements are linked. Even during the financial crisis they tend to stay in close proximity to their sector.

TABLE 3.1: 10-period volatility correlation of the Insurance sector.

	RSA	AV/	PRU	SL/	Sector Avg.
RSA	1.00	0.72	0.76	0.80	0.84
AV/	0.72	1.00	0.90	0.86	0.95
PRU	0.76	0.90	1.00	0.86	0.94
SL/	0.80	0.86	0.86	1.00	0.93
Sector Avg.	0.84	0.95	0.94	0.93	1.00

Specimen of the insurance sector in the FTSE 100. We present the largest companies in the sector. The correlation between the 10-period volatility measured between stocks and the sector average through standard deviation is high, also depicted in Figure 3.12.

For the long only model, the null hypothesis is set in reference to the Kurskal-Wallis test. Stocks from the same sector must have similar variance, coming from the same distribution. The null hypothesis is that stocks do not have similar variance and if the they pass the test the null hypothesis is rejected. We will use these tests in Chapter 5, where we build quantitative investment models.

3.1.5 Summary

In this chapter we presented several statistical tests, which were used to identify structure in data, as well as to validate some expectations we have from financial data as presented in the previous chapter. These tests help us get a better understanding of our data, as well as validate the path we take to build some of our models, since some of the models would not work if certain characteristics in the data did not exist. To give context to our statistical tests we grouped them by the quantitative investment model they are relevant to, as well set out relevant null hypothesis to reject. These tests not only help find structure in data, but also help identify the parameters of the eventual model. For example the Orenstein-Uhlenbeck model will be used for market neutral pairs QIMs, while regression analysis will be applied to nearly all of the QIMs where we need to establish relationships between stocks in the same sector, as well markets in the same time zone.

These statistical models play an important role in helping us build investment models that have a high probability of success, which we identified as one of the important requirements of good portfolio management in Chapter 1. In the next chapter we introduce the framework within which all our models will work. We will discuss how all the models operate and interact with each other, as well the stock market within this framework.

Chapter 4

A Framework of Quantitative Investment Models

In the previous chapter we discussed some statistical tests and models that will play a key role in building our QIMs. In this chapter we will begin by introducing the broad framework within which our models will operate. We will then present how the models interact within this framework and the market as well as other models. We will also introduce the four quantitative investment models (QIMs) that we will use to capture different features in FTS data.

4.1 Model Framework and Approach

As discussed in Chapter 1, the aim of our research to combine models that capture different aspects of patterns and structure in data found at differing time horizons: long-run models to capture momentum and short-run models to capture mean reversion. In this chapter we present the broad framework within which our models operate throughout the rest of the thesis. We will then present role of these models within this framework and the market, as well as other models. Our framework can be viewed as a two-tier system: Tier one is where the models are designed to focus on a particular aspect of FTS, such as momentum. These models interact with the market and take trading decisions. Tier two has the model that allocates capital to the models in Tier one, essentially building a portfolio of quantitative models.

For our research we will implement four quantitative investment models and five meta models, where the meta models allocate capital to the four quantitative investment models. We choose these four models as they capture momentum and mean reversion well in equities as an asset class, at different time horizons. Even though these models operate on the same data or members of the same data set, we expect them to have low correlations. Each of these four quantitative models are designed to focus on their investment approach, without interfering in the working of the other models.

4.1.1 Model Framework

In this section we present the framework within which our models operate. To this end, Figure 4.1, presents the framework diagrammatically, which also shows the flow of information within the framework.

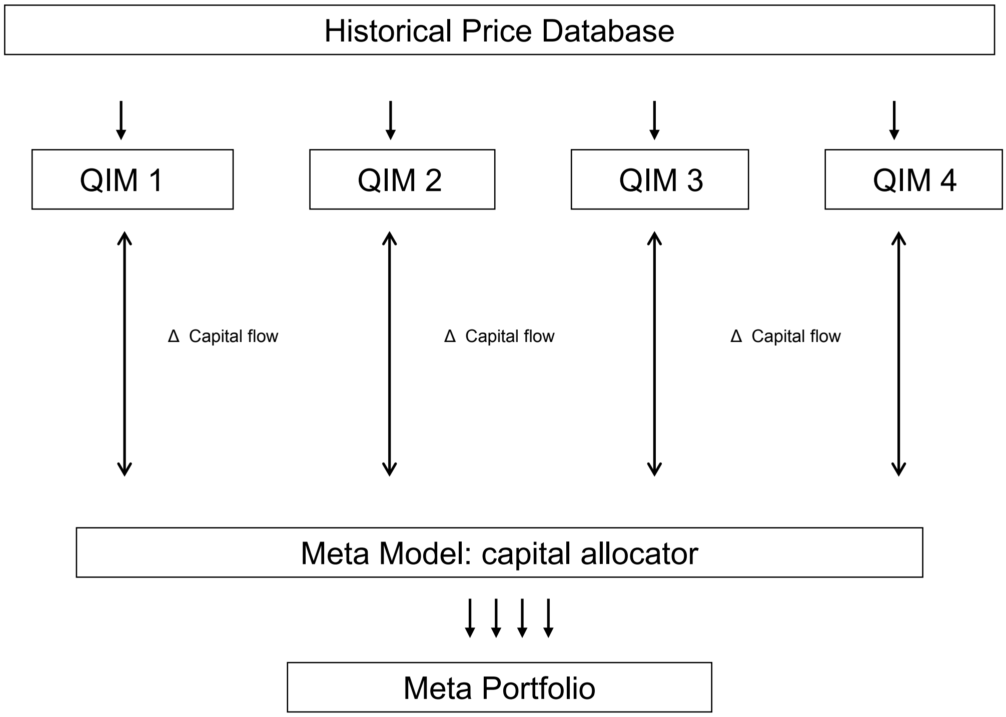


FIGURE 4.1: Flow diagram for financial data, underlying investment models and meta portfolio. This figure shows the broad framework within which our models will interact.

In more detail the framework consists of the following components:

- **Historical price database:** The database contains FTS of daily prices the open: high, low and close (OHLC) of each and every trading day in the data set. It supports the QIMs. This database also supports the analytics of all the models, and the valuation of portfolios on a daily basis.
- **The quantitative investment models:** Quantitative models are simply investment methodologies, where the process is driven by an algorithm (Section 2.3). Here, each QIM follows its own distinct methodology of investing its assigned capital without interfering with other QIMs. We distinguish between quantitative models and models that are quantitative as well as systematic. As mentioned in

Section 2.3, by systematic and quantitative we mean that the initiation and implementation of an investment decision is entirely controlled by an algorithm, i.e. there is no human intervention. For our research we introduce four quantitative models. All these models operate using daily data from the database, using the capital allocated to them from the meta model. We give a more detailed explanation of the individual models in Section 4.1.2.

- **Meta model:** The meta model is a dynamic allocator of investment capital to the four models. Although the data and analytics are updated daily, the meta model is designed to make capital allocation decisions at the end of every month. For example, if the meta model takes a decision to move investment budget from QIM 1 to QIM 4 and QIM 1 happens to be fully invested, in such a scenario, QIM 1 will liquidate part of its portfolio to make cash available to be moved to QIM 4. The liquidation to make capital available is done universally across all holdings so as not to change the distribution of the portfolio. This will be the standard process for the inflows and outflows of all our quantitative investment models.

- **Meta portfolio:** The meta portfolio represents the value of holdings in portfolio which is the outcome of capital allocations made to QIM 1, QIM 2, QIM 3, and QIM 4. The performance of the meta portfolio is the weighted average of the performance of the four models and any unallocated investment capital is invested in risk free government bonds. Here, the **weights** refer to the percentage of capital assigned to each of the four models.

4.1.2 Quantitative Investment Models and Interactions

We will now discuss the role of our QIMs and how they fit within the framework that we presented in Figure 4.1. Specifically, the role of each QIM is to focus on a particular pattern identified, on a certain time frame, in its distinct data set, where the pattern has been identified. We explain this further with Figure 4.2. As we can see, the two key inputs into a QIM are i) data and ii) investment capital. Both data and investment capital are essential for the QIM to operate. The data is used by the computation engine to identify investment opportunities in the market. Once these investment opportunities have been identified the QIM needs capital so that it can interact with the market to make investments. Once these investments have been made the QIM has begun the process of building a portfolio. The portfolio is the ultimate objective of the QIM. The portfolio of each QIM becomes part of the meta portfolio as shown in Figure 4.1.

In a live market setting, the QIM would work as follows: At the start of the day the QIM connects to the price database and updates its data to the latest observations. Once the latest data is obtained, the models checks for gaps and errors in the data. With data now checked, the QIM updates the values of its portfolio which it has from the previous day. With the portfolio updated with the latest prices, the QIM runs its algorithm and generates actions that it needs to take today depending on the capital it has available. Now that the QIM is ready for the day, the QIM connects to its assigned account in the server of the stock exchange. Having established its connection to the stock exchange the QIM is now ready to send buy or sell order to the exchange.

Change in allocation of investment capital to each QIM is done on a monthly basis. This is for several reasons. Firstly, higher frequency changes have a lot more noise so the estimates are not good. Secondly, the QIMs need time to generate some returns before one reassigns capital. Finally, from a practical perspective, in real life if people reallocated capital to QIMs everyday, they would disrupt the portfolio too often for it actually settle in and generate some returns.

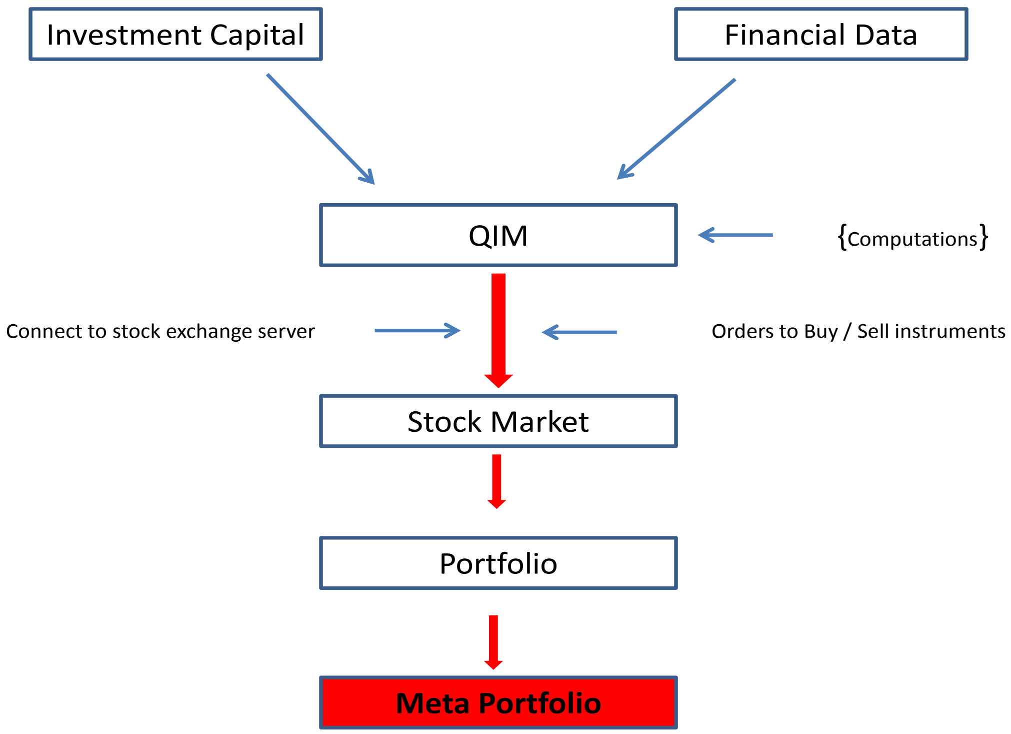


FIGURE 4.2: Interaction and workings of a quantitative investment model that, builds a portfolio targeting a distinct pattern on a particular time horizon, which eventually becomes part of the meta portfolio.

Within this framework we essentially have five models, four QIMs and one meta model completing the framework. The QIMs are the following:

- Our first model is a long only model (see Section 5.1.3). This model focuses on buying stocks that are lagging behind their peer group, where the peer group is defined by sectors. In each sector grouping we buy the laggards. We implement this model on FTSE 100 stocks.

- Our second model is a market neutral model for stocks, using the pairs approach (see Section 5.1.2). This model focuses on mean reversion between relationships that have temporarily dislocated from their historical path. This model is implemented by choosing the most suitable stock pairs from the universe of 250 of the largest European stocks, denominated in Euros.

- Our third model is a momentum model (see Section 5.1.1). This model focuses on identifying long-run trends in the market and takes positions by buying or selling relevant assets to profit from these trends. This model focuses on global equity indices that have active futures markets.

- Our fourth model is a market neutral model (MN) for global equity indices (see Section 5.1.2). This model focuses on mean-reversion between relationships that have temporarily dislocated from their historical path. We implement this model at the index level. This model also focuses on global equity indices.

- Our meta model(s) will allocate capital to the four quantitative investment models. The resultant outcome is a portfolio of these four models, which we shall refer to as the meta portfolio. We discuss them in more detail in the next section.

As we fill in the framework presented in Figure 4.1 with our four QIMs our working framework will look like Figure 4.3, taking us to the meta portfolio. In Chapter 5 we will build these models and go into more detail about how they operate.

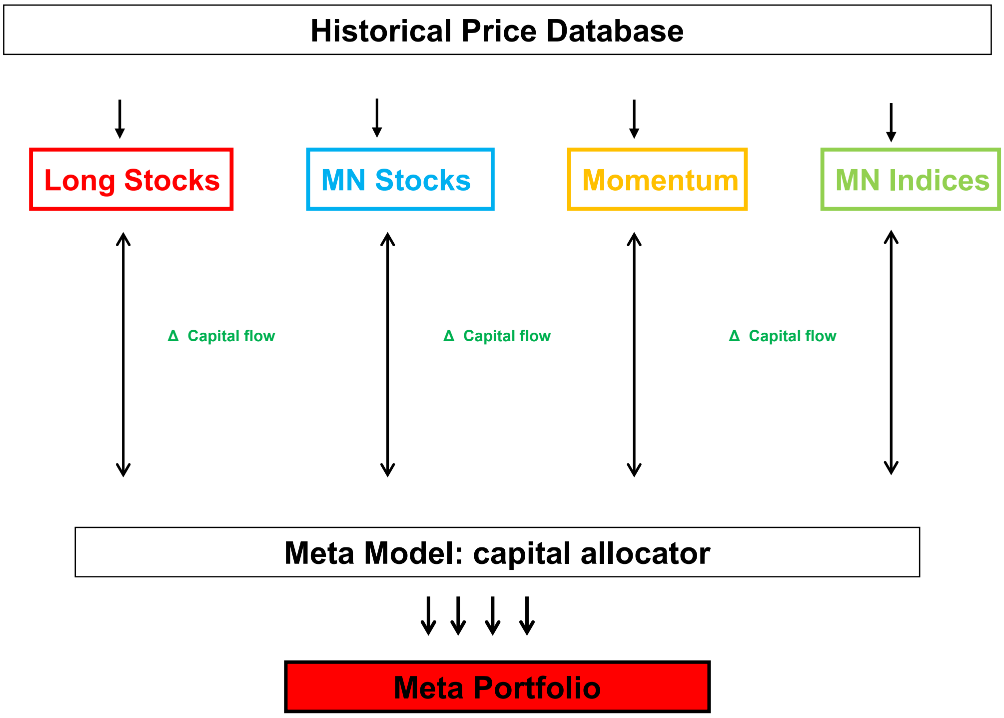


FIGURE 4.3: Flow diagram of the complete framework. This includes the four QIMs i) Long stocks, ii) MN stocks, iii) Momentum, and iv) MN indices and the meta portfolio.

4.2 Summary

In this chapter we introduced the framework within which we will build and operate our models through the rest of the thesis. Towards the end of the thesis in Chapter 8 we will be able to see why this two-tier approach works well, especially when we compare the performance of the QIMs in the tier that interacts with the market and the meta models. This two-tier framework helps break down the challenge of good portfolio management as discussed in Chapter 1, namely investment opportunity with a high probability of success (Tier 1) and correct investment size and risk control (Tier 2). In the following chapters we will build these models and populate this framework.

In the next chapter we will present the QIMs that focus on particular patterns in the data and interact with the market. In Chapter 6 we will present the meta models, which constitute the second tier of the framework (see Figure 4.3), leading to the meta portfolio.

Chapter 5

Constructing Quantitative Investment Models

In the previous chapter we presented the framework within which all our models will operate, (see Figure 4.3). The framework has two-tier. The first-tier handles the data, analytics, and interacts with the market. The second tier has the meta model that allocates capital to the first tier leading to the meta portfolio. The first tier has four QIMs. In this chapter we will build these QIMs. These QIMs are targeting momentum and mean-reversion found in price time series data.

In Section 5.1 we will start with a detailed description of our QIMs. We present the reasoning behind the models and how we use the statistical tests and parametrise the models. We follow up the QIMs with our meta models in Chapter 6. We have built six meta models and they are compared with three benchmarks when we present the results in the final chapter. The meta models are our key contribution.

5.1 Quantitative Investment Models

In this section we present the QIMs that interact with the market in more detail. We show how they are structured and calibrated in order to capture structure in data. The output of these models will become part of the meta portfolio. After presenting the QIMs we present our meta models, which allocate investment capital to the QIMs.

5.1.1 Momentum Model for Global Equity Indices

Our first model is the momentum model. We shall use this model for the global equity index futures market. Here a futures contract is an agreement between two parties to buy or sell an asset at a certain time in the future for a certain price (Hull, 2011). The objective of the momentum model is to identify long-run directional move up and down, i.e: a continuation of returns with the same sign for equity index futures market, and to position the portfolio to gain from these trends. The model is built using a simple moving average (see Section 2.3 for details) and Kalman Filter (Section 2.5). The moving average is calculated on a fixed length rolling window.

To build this model we need to estimate the window size of the moving average. We use regression analysis performed by Moskowitz et al. (2011) discussed in Section 3.1.3, where the null hypothesis is also set in context to the regression analysis. Moskowitz et al. use regression analysis to find predictability in returns, based on lagged returns. The lag length for which we have positive coefficients with corresponding t-statistics serves as the initial reference point to identify the lag length. Once we have this lag length we search for the best window around this period in daily data, so that all equity index futures are profitable and the momentum portfolio achieves a high Sharpe ratio.

All QIMs interact with the market at daily frequency using closing prices from the market (see Section 4.1.1). To reduce the amount of noise and jumps in financial data at daily frequency and make a estimate reliable we will use a Kalman Filter. The forecast of the Kalman Filter will be used to construct the moving average.

Our Kalman Filter is a simple scalar model: i.e. there is just one variable, the stock price. Our approach is that today's closing price (at time t) is the the best estimate for tomorrow's price (at time $t + 1$), with some error. Hence our state model coefficient is set to unity (1) and the variance is estimated using polynomial fit on the first 50 data points. The order of the polynomial is derived from using half the degrees of freedom. The variance is then set to 1/10 of its value. Hence the only parameter that is set is the variance in the model $W_t = W$ and the variance in the observation $V_t = V$. The Kalman Filter stores the estimates of θ_t called $\theta_{t|t}$and estimated variance as $P_{t|t}$ (see Algorithm 1).

The model operates by using the simple moving average and price data. We step through the model in Algorithm 2. After initialising the Kalman Filter and updating the moving average, we check if the price of a futures is above the moving average. If that is the case we check if whether we already have a position open or not. If not, we buy futures (long) and open position and update the portfolio. If the price is below the moving average and if we do not have a position open, we open a new position by selling (short) and updating the portfolio.

Trades are done at the close of trading. We assume we can trade at the closing price and all trades are adjusted for investment capital at close. The capital allocation to each trade in the momentum model is equally divided. They are fixed as the inverse of the number of markets i.e: if there are 20 markets to invest in, then the investment capital is fixed at 1/20 or 5%. We allocate investment capital in equal amounts so that we can compare our model to the equally weighted buy and hold (BAH) approach for the investment universe to show that pour model adds value. The model is always invested in the market and has no stop loss. Once we have the performance results we convert the performance data to monthly data to used for the meta models.

Algorithm 1 *KALMAN FILTER Function*: *Input: price data,* W *= model variance,* V *= observation variance. Output:* $\theta_{t|t}$.

function KALMAN FILTER(z_t)
  if t = 0 **then**
    initialise $\theta_{0|0} \leftarrow z_0$
    initialise $P_{0|0} \leftarrow 1$
  end if

  $\hat{\theta}_{t|t-1} \leftarrow \hat{\theta}_{t-1|t-1}$,
  $P_{t|t-1} \leftarrow P_{t-1|t-1} + W_t$,
  $y_t \leftarrow z_t - \hat{\theta}_{t|t-1}$,
  $S_t \leftarrow P_{t|t-1} + V_t$,
  $K_t \leftarrow P_{t|t-1} S_t^{-1}$,
  $\hat{\theta}_{t|t} \leftarrow \hat{\theta}_{t|t-1} + K_t r_t$,
  $P_{t|t} \leftarrow (I - K_t H_t) P_{t|t-1}$
  store $\theta_{t|t}$ *and* $P_{t|t}$
  return $\theta_{t|t}$
end function

5.1.2 Market Neutral for Global Equity Indices and EU Stocks

The market neutral model will be used for the global equity index futures and EU stocks. This model is structured in a 1 unit vs. 1 unit fashion. This is the dollar neutral pairs model we discussed in Section 2.3. The objective of this model is to capture mean reversion in “related time series or pairs” that have dislocated from the their long-term mean. The pairs by design are set up as the difference between log of two times series, e.g. two stocks such as $\log(StockA) - \log(StockB)$, or two indices such as $\log(IndexX) - \log(IndexY)$, leading to a third time series which we call the *spread*, as shown in Figure 3.4.

To identify relevant stocks and futures to construct a spread, this model carries some additional structure for both global equity index futures as well as EU stock pairs. For global index futures we create geographic time zones to pair indices. The reason we do this is that markets that operate in the same time zone also process news and react together. Markets in time zones that come much later have more time to process new information and have a slightly different reaction in their markets, again impacting correlation among them. In the case of EU stock pairs, the structure is in the form of sectors and sub-sectors. Stock pairs are made from stocks that come from the same sector or sub-sector, making the model both dollar and sector neutral. We will discuss this structure in more detail in Section 8.2.

Algorithm 2 *Momentum*: Input x: Log of index futures, Output: Momentum model portfolio.

```
# READ daily futures prices as x_(n,t), where t is time stamp and n is the number of
futures markets and x ∈ ℝ.
# rebalance() is a function that takes the current portfolio, computes the mean in-
vestment in the futures market and buys and sells so that capital is equally divided
among futures markets.
# w = moving average window size
# Π_t = (Π_(1,t), Π_(2,t)...Π_(n,t)), where Π_(i,t) = ⟨C_(i,t), P_(i,t)⟩
# C_(i,t) is capital invested in futures contract i at time t.
# P_t = (P_(1,t), P_(2,t)...P_(n,t)) position in portfolio.
# P_(i,t) ∈ {Buy, Sell} position of futures i at time t
# KALMAN FILTER is the Kalman Filter Function as described in Algorithm 1.

for i = 1 to n do
  C_(i,0) ← investment/n
  P_(i,0) ← Buy
end for

for t = 1 to T do
      C_t ← rebalance(C_(t−1))
      P_t ← P_(t−1)
   for i = 1 to n  do
         x_(i,t) ← get current price
         y_(i,t) ← KALMAN FILTER(x_(i,t)) # Kalman filter prediction
         z_(i,t) ← moving average (y_(i,t))          (1/w) Σ_{t−w}^{t} y_(i,t)
      if x_(i,t) > z_(i,t) then
         if P_(i,t) != Buy then
               P_(i,t) ← Buy
         end if
      else
         if P_(i,t) != Sell then
               P_(i,t) ← Sell
         end if
      end if
   end for
end for
```

Once the initial pairs have been identified we begin with a set of statistical tests that we discussed in Section 3.1.2. The tests serve as an initial indicator as to whether relationships are stable. The regression analysis shows that a relationship exists between two time series; the residue from the regression shows us that the relationship can be potentially mean-reverting. We want the residual crossing the mean as frequently in the in-sample data set. The Runs test gives us an idea of how much time it takes for the residual to revert, giving a sense of autocorrelation in the residual. We also check for mean reversion in the spread using the VRT, which gives an indication on the window size on which to focus of the final model. Most importantly we check for mean-reversion in the spreads using VRP, which tells us about mean-reversion as per the null hypothesis (see Section 3.1.2). We use the Ornstein-Uhlenbeck model to model the spread. The final threshold is the success of a spread to be included in the portfolio.

Tests serve an important role in guiding us to identify key characteristics in the data, the model and window size. Building and calibrating the model is an iterative process, since we need to find the best window size where we get the highest profitability in our trades as well as the portfolio. We begin with a standardized window size of 10 days and go upto 50 days in increments of 5 days, where each increment is equal to 1 working week. At each window size we need to identify the threshold for lambda calculated using the Ornstein-Uhlenbeck model, which gives us the speed of mean reversion and the standard deviation measure that gives us the best return.

The input in Algorithm 3 is the spread chosen for their profitability. For each time step for each spread, we compute the mean, standard deviation and lambda. If lambda is less than its threshold and standard deviation is greater than its threshold, check if a position is already open. If not, then open a position by selling the spread and update the portfolio. If lambda is greater than its threshold and standard deviation is greater than its threshold, check if position is already open. If not then open a position by buying the spread and update the portfolio. If the spread is equal to the mean and position is open, then close position and update portfolio.

The model operates with three important thresholds, sequentially lambda, standard deviation and mean. Trades are executed at the close of trading, and we assume we can trade at the closing price and all trades are adjusted for investment capital at close. We also assume that the trade size is small and there is no market impact. The investment capital for each trade in this model is equally divided; fixed as the inverse of the number of open pairs i.e. if there are 20 pairs to invest in, then the investment capital is fixed at 1/20 or 5%.

Algorithm 3 *Market Neutral*: Input S: Log spread of two paired time series, Output: market neutral model portfolio.

READ daily spread values as $S_{(n,t)}$, where t is time stamp and n is the number of spreads and $S_{(n,t)} \in \mathbb{R}$.
rebalance() is a function that takes the current portfolio, computes the mean investment in the futures and stock market and buys and sells so that capital is equally divided among futures markets.
release() is a function that takes the current portfolio and sells enough open position across all holdings to release a proportion of $\frac{1}{n+1}$ of the capital so that it can open a new position.
realise() is a function that closes open position i which is then rebalanced.
$\Pi_t = (\Pi_{1,t}, \Pi_{2,t}...\Pi_{n,t})$, where $\Pi_t = \langle C_{(i,t)}, P_{(i,t)} \rangle$
$C_{(i,t)}$ is capital invested in futures contract i at time t.
$C_{(t)}$ is capital value of the portfolio at t.
$P_t = (P_{(1,t)}, P_{(2,t)}...P_{(n,t)})$ position in portfolio.
$P_{(i,t)} \in \{Buy, Sell, Not\_Own\}$ position of futures i at time t
OU is a function which uses the Ornstein-Uhlenbeck model, described in equation Equation 3.3 to compute λ.
mean is a function which calculates the mean of the data.
std is a function which calculates the standard deviation of the data.
w is the window size for calculations.
λ = speed of mean reversion,
γ = mean reversion threshold,
μ = mean of the spread,
ϕ = standard deviation threshold,
σ = standard deviation of the log spread.

```
for t = 1 to T do
    C_t ← rebalance(C_{t−1})
   for i = 1 to n do
     Initialise OU
     S_(i,t) ← get current spread
     λ_(i,t) ← OU(S_(i,t), S_(i,t−1), S_(i,t−2)...S_(i,t−w)) # compute lambda
     μ_(i,t) ← mean(S_(i,t), S_(i,t−1), S_(i,t−2)...S_(i,t−w)) # compute mean
     σ_(i,t) ← std(S_(i,t), S_(i,t−1), S_(i,t−2)...S_(i,t−w)) # compute standard deviation.
      if λ_(i,t) < −γ & σ_(i,t) > φ then
         if P_(i,t)! = Buy then
              C_t ← release(C_t)
                P_(i,t) ← Buy
         end if
      else
         if λ_(i,t) > γ & σ_(i,t) > φ then
             if P_(i,t)! = Sell then
                C_t ← release(C_t)
                  P_(i,t) ← Sell
             end if
         end if
      end if
      if S_(i,t) = μ_(i,t) then
           P_(i,t) = Not_Own
             C_t ← realise(C_t)
      end if
   end for
end for
```

5.1.3 Long Only for Stocks

In the long only model we only "buy" stocks and maintain a constant exposure to the market at all times. The objective of this model is to buy stocks that have lagged their peer group in performance, as according to financial theory, stocks from the same sector are supposed to move in tandem with their peer group. We use both open and close prices for this model, capturing a little more structure in data at higher frequency. To build this model we start by classifying stocks into their sectors. Once we have created sector groups, we create a sector index, which is the average of stocks in that sector. Once we have the data and sector index classified properly, we run the tests that we discussed in Section 3.1.4. Regression analysis checks to see if the stocks are closely related. Most importantly, the Kurskal-Walis test tells us that the returns came from same distribution; this is also the null hypothesis and short run volatility should also be correlated.

The long only model is a constant rebalancing model; i.e. we rebalance the portfolio at every time step and maintain equal capital allocation to all positions in each sector. Trades are done at the open and close of trading; we assume we can get the opening and closing price and all trades are adjusted for investment capital at trading time. We maintain constant exposure in the model, i.e. we are always invested; we do not have any stop loss on this strategy. We pursue the following process:

At every time step, for each sector, we rank the stocks by their returns in descending order. We then divide the number of stocks by two dividing the sector constituents into two halves. The top half is better performing than the bottom half. We round the result of the division down to the nearest whole number if the number is not divisible by two, identifying the stocks to buy. At the next time step, we buy the stocks that are in the bottom half of the rank. Then the model updates the portfolio, profit and loss and the capital to assigned to all the trades at the next time step.

Specifically, we buy the worst performer with the expectation that they will catch up with the other, better performing stocks. For example in the current period we invest (buy) the data series that have gone up the least. For example, if we have 5 data points in a group, namely A, B, C, D and E, and after ranking, E is ranked 1^{st}, C is ranked 2^{nd}, B is ranked 3^{rd}, D is ranked 4^{th} and A is ranked 5^{th}. So for the next period we would invest (buy) D and A as they are the weakest performers.

Now that we have the returns of all the QIMs, we take their returns at monthly frequency to be used in the meta models, as we change our capital allocation in the meta model on a monthly basis. The reason we do this is because it is impractical to change allocation at a faster rate not only in real life, but also because data at higher frequency has higher degree of noise, making estimation harder even for simple statistics such as mean and variance.

Algorithm 4 *Long Only*: Input $\mathcal{E}$: Stock returns, Output: Long Only portfolio.

READ stock returns prices as $\mathcal{E}_{(n,t)}$, where t is time stamp and n is the number or stocks and $\mathcal{E}_{(n,t)} \in \mathbb{R}$.
$rebalance()$ is a function that takes the current portfolio, computes the mean investment in the futures market and buys and sells so that capital is equally divided among futures markets.
$\Pi_t = (\Pi_{1,t}, \Pi_{2,t} ... \Pi_{n,t})$, where $\Pi_t = \langle C_{(i,t)}, P_{(i,t)} \rangle$
$C_{(i,t)}$ is capital invested in futures contract i at time t.
$C_{(t)}$ is capital value of the portfolio at t.
$P_t = (P_{(1,t)}, P_{(2,t)} ... P_{(n,t)})$ position in portfolio.
$P_{(i,t)} \in \{Own, Not\_Own\}$ position of futures i at time t
Let $\mathcal{S}$ be set of sectors.
Let s is a sector $\subset \mathcal{S}$.
Let $\mathcal{E}$ be set of equities in sector $\mathcal{S}$.
Let e be single equity.
Let $\mathcal{L}$ be set of laggard stocks in a sector.
$laggard()$ is a function that returns $\lfloor |\mathcal{E}|/2 \rfloor$, the worst performing stocks at current time.

for i = 1 to n **do**
  $C_{(i,0)} \leftarrow investment/n$
end for
for $t = 1$ to T **do**
  $C_{(t)} \leftarrow rebalance(C_{(t-1)})$
  for $s \in \mathcal{S}$ **do**
    $\mathcal{L}_{(s,t)} \leftarrow laggard(\mathcal{E}_{(s,t)})$
    for $e \in \mathcal{E}_s$ **do**
      if $P_{(e,t)} = Own$ and $e \notin \mathcal{L}_{(s,t)}$ **then**
        $Sell(e)$
      end if
    end for
    for $e \in \mathcal{L}_s$ **do**
      if $P_{(e,t)} = Not\_Own$ **then**
        $Buy(e)$
      end if
    end for
  end for
end for

5.2 Summary

In this chapter we presented the QIMs that will capture patterns in data at different time horizons. We presented four QIMs,

- Momentum model,
- Market neutral model for i) equity index pairs and ii) stock pairs and
- Long only model.

The momentum and the market neutral models used closing prices but the long only model used both opening and closing prices to capture more structure in data. These models are not new; many variations of these models have existed. However, they serve our purpose as they operate on different time horizons, capturing patterns at different time horizons. These models help us address the challenge of investment opportunities that have high probability of success as discussed in Chapter 1 as essential for good portfolio management. In the next chapter we present the meta models which contain our main contributions and address the challenge of correct investment size and risk control, the other two important components of good portfolio management. These QIMs will get their investment capital from these meta models, giving us a portfolio of QIMs and leading to the meta portfolio.

Chapter 6

Constructing Meta Models

In the previous chapter we presented the QIMs that will manage investments in the markets. In this chapter we build a set of capital allocation models which allocate investment capital to our four QIMs. As discussed in Chapter 4 we call this the meta model, which gives us the meta portfolio. In doing so we aim to construct portfolios such that they have the highest risk adjusted return, as measured by the Sharpe ratio. All the models work within the framework presented in Chapter 4, Figure 4.3.

In Section 6.2 we will start with a detailed description of our meta models. We present the reasoning behind the models how we use statistical tests to parametrise the models. We have built five meta models and they are compared with four benchmarks in the final chapter. The meta models are our key contribution.

6.1 Functions

In this chapter we will present several meta models. These meta models will have representation in pseudo code. We will create some functions which will be called into the algorithms. We present these functions below.

6.1.1 Fractional Kelly Function

This function (Algorithm 5) calculates the Kelly value for input data, which in our case is the data from the QIMs, using the formula presented in Section 2.2.4, Equation 2.9. Where the calculation is negative we replace the negative value with a zero. This is because we are only interested in models where there is a positive expectation. When the expectation for a certain model is negative we simply do not invest in that model and allocate capital to profitable models for that time step. We then calculate Fractional

Kelly. We do this by normalising the positive Kelly estimates. This gives us Fractional Kelly weights.

Algorithm 5 *KELLY Function*: *Input* $\mathbf{X}$ = QIM Returns data and r as risk-free rate, *Output* = K, Fractional Kelly weights.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
r is the risk-free rate.
$mean()$ is a function which calculates the mean of the data.
$var()$ is a function which calculates the variance of the data.
$K = (K_{(1)}...K_{(n)})$ are fractional Kelly weights.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

$sum \leftarrow 0$
for $t = 1$ to T **do**
  for $i = 1$ to n **do**
    $\mu_{(i)} \leftarrow mean(\mathbf{x}_{(i,1:t)})$
    $\sigma^2_{(i)} \leftarrow var(\mathbf{x}_{(i,1:t)})$
    $K_{(i)} \leftarrow \frac{\mu_{(i)} - r}{\sigma^2_{(i)}}$ ▷ (Equation 2.9)
    $sum \leftarrow = sum + K_{(i)}$
  end for
end for
for $i = 1$ to n **do**
    $K_{(i)} \leftarrow \frac{K_{(i)}}{sum}$
end for
return K

6.1.2 Median Kelly Function

The median Kelly function shown in Algorithm 6 calculates the Kelly value for input data, which in our case is the data from the QIMs. The median Kelly function operates exactly like the Fractional Kelly function except that the Kelly calculations are done using the median of the data signified by $\tilde{x}$ to estimate Kelly instead of the mean as previously shown in Equation 2.9, all other parameters being the same. We then estimate the allocation percentage as we do for Fractional Kelly (Section 6.2.1.3) using Equation 6.1.

$$f^* = \frac{\tilde{x} - r}{\sigma^2}. \tag{6.1}$$

Algorithm 6 *MEDIAN KELLY Function*: *Input* $\mathbf{X}$ = QIM Returns data and r risk-free rate, *Output* = MK, Median Kelly weights.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
r is the risk-free rate.
$mean()$ is a function which calculates the mean of the data.
$var()$ is a function which calculates the variance of the data.
$MK = (MK_{(1)}...MK_{(n)})$ are median Kelly weights.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

$sum \leftarrow 0$
for $t = 1$ to T **do**
  for $i = 1$ to n **do**
    $\tilde{x} \leftarrow median(\mathbf{x}_{(i,1:t)})$
    $\sigma^2_{(i)} \leftarrow var(\mathbf{x}_{(i,1:t)})$
    $MK_{(i)} \leftarrow \frac{\tilde{x}-r}{\sigma^2}$ ▷ (Equation 6.1)
    $sum \leftarrow = sum + MK_{(i)}$
  end for
end for
for $i = 1$ to n **do**
    $MK_{(i)} \leftarrow \frac{MK_{(i)}}{sum}$
end for
return MK

6.1.3 Performance Curve Function

The performance curve function as shown on Algorithm 7 represents the performance of every QIM based on an initial investment of 100 units using this formula

$$PC_{(t-1,m)} * (1 + (QIM_{(t,m)}). \tag{6.2}$$

Specifically 6.2 represents the compounded growth of the initial investment of 100. Here QIM represents the return on the model and PC the price. We will use the performance curve function in four models. Kelly with Kalman Filter, Median Kelly with Kalman Filter, Kelly with Moving Average and Median Kelly with Moving Average.

6.1.4 Kalman Filter Function

This function uses the data from the QIMs (Algorithm 8). The function attempts to forecast whether the QIM return in the next period is positive or negative. The input

Algorithm 7 *PERFORMANCE CURVE Function*: *Input* = QIM Returns, *Output* = *price*, Performance curve of QIM.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$price_{(i,t)}$ is the value of the model indexed to 100

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

$price_{(1:n,0)} \leftarrow 100$
for $t = 1$ to T **do**
  for $i = 1$ to n **do**
    $price_{(i,t)} \leftarrow price_{(i,t-1)} \times (1 + \mathbf{x}_{(i,t)})$ ▷ Equation 6.2
  end for
end for
return *price*

to this model is the monthly performance of the QIMs. The function checks whether the forecast for $(i+1)$ is positive or negative when compared to the previous time step. When the Kalman Filter forecast is negative, the signal is converted to 0 and when the forecast is positive the signal is 1. The reason we change the forecast to binary data is to adjust the Kelly in the upcoming meta models. To use the Kalman Filter we need the QIM data. We convert the QIM data into its multiplicative form as shown in Algorithm 7.

6.2 Meta Models

We now present our main contribution, the meta models that will be used to allocate investment capital to our QIMs. In this section we present several models using the Kelly Criterion. We will present Median Kelly, Kelly with Kalman Filter, Median Kelly with Kalman Filter, Kelly with Moving Average and Median Kelly with Moving Average. Our benchmark will be Equally Weighted, MVO, Fractional Kelly and Optimal Kelly. The key challenge with Kelly portfolios is maximising the Sharpe ratio. With a range of models we endeavour to address this challenge. None of our models use any leverage, hence all models have the same amount of initial capital. All Kelly models are fractional in nature i.e. even if Kelly estimates suggest that one should use leverage, the estimates are trimmed to be fractional.

Algorithm 8 *BINARY KALMAN FILTER Function*: *Input* = z_t, Performance curve of QIM, W is model variance, V = observation variance *Output* = 0 or 1, Signal for QIM.

Binary Kalman Filter forecast for t+1 converted to binary signal
if t = 0 **then**
    initialise $\theta_{0|0} \leftarrow z_0$
    initialise $P_{0|0} \leftarrow 1$
end if

for $t = 1$ to T **do**
    $\hat{\theta}_{t|t-1} \leftarrow \hat{\theta}_{t-1|t-1}$,
    $P_{t|t-1} \leftarrow P_{t-1|t-1} + W$,
    $y_t \leftarrow z_t - \hat{\theta}_{t|t-1}$,
    $S_t \leftarrow P_{t|t-1} + V$,
    $K_t \leftarrow P_{t|t-1} S_t^{-1}$,
    $\hat{\theta}_{t|t} \leftarrow \hat{\theta}_{t|t-1} + K_t r_t$,
    $P_{t|t} \leftarrow (I - K_t H_t) P_{t|t-1}$
    store $\theta_{t|t}$ *and* $P_{t|t}$
    return $\theta_{t|t}$

    if $\theta_{t|t} \leq \theta_{t-1}$ **then**
        return 0
    else return 1
    end if
end for

6.2.1 Meta Model Benchmarks

We present a few meta models but to compare them we chose four benchmarks, a) Equally Weighted, b) Mean Variance framework, c) Fractional Kelly and d) Optimal Kelly. Equally weighted is chosen as it is a model free approach without assumptions. The Mean Variance framework is chosen, as it is not only the most well known and popular approach, it is also the toughest approach to beat. The Optimal and Fractional Kelly approaches use Kelly for portfolio construction, which can be used to compare our meta models.

6.2.1.1 Equally Weighted

Our first benchmark is the Equally Weighted (EW) method. This is chosen as it is a model free approach which comes without assumptions. Although it seems easy, in some instances it can be hard to beat in terms of returns. The EW method is straightforward in terms of computation. We outline the approach in Algorithm 9. *At each time step the QIMs are given equal weight.* Once we have the returns from the previous period calculated, we can calculate the new investment capital available and then make the adjustments to equally distribute it. For example if we have four QIMs then we will have a 25% allocation to each of the QIMs at all times.

Algorithm 9 *Equally Weighted*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\Pi_{(t)} = (\Pi_{(1,t)}, \Pi_{(2,t)} ... \Pi_{(n,t)})$.
withdraw_capital() is a function which withdraws capital from a QIM is it has more capital than the average allocation to all QIMs.
allocate_capital() is a function which allocates capital to a QIM is it has less capital than the average allocation to all QIMs.

$$\mathbf{\Pi} = \begin{bmatrix} \Pi_{11} & \Pi_{12} & \dots & \Pi_{tn} \\ \Pi_{21} & \Pi_{22} & \dots & \Pi_{tn} \\ \dots\dots & & & \\ \Pi_{T1} & \Pi_{T2} & \dots & \Pi_{Tn} \end{bmatrix}$$

for $t = 1$ to T **do**
  $average \leftarrow \frac{\mathcal{P}_{(t-1)}}{n}$
  for $i = 1$ to n **do**
    if $\Pi_{(i,t)} > average$ **then**
      $withdraw\_capital(\Pi_{(i,t)} - average)$
    else if $\Pi_{(i,t)} < average$ **then**
      $allocate\_capital(average - \Pi_{(i,t)})$
    end if
  end for
  $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \Pi_{(i,t)}$
end for

6.2.1.2 Mean Variance Optimisation

Our second benchmark will be the mean-variance framework (MVO) framework introduced by Harry Markowitz in 1952 and 1959. MVO remains a *tour de force* in the field of portfolio construction. Even though its has had it fair share of criticisms, it still remains hard to beat. Markowitz's original MVO model was static. However in Section 2.2.1 we discussed that dynamic models have given better results. Hence we will use the dynamic approach, where we update our estimates as new data becomes available at the next time step.

In Algorithm 10, QIM returns are the input into the model from which we estimate average returns and covariance matrix at each time step. The objective here is to identify the optimal weights identified by the highest Sharpe ratio. We take the allocation weights from the optimal portfolio. The optimal weights are then used for the portfolio in the next time step $(t + 1)$ giving us the portfolio return. Then the model updates the portfolio, profit and loss and the capital assigned to all the trades at the next time step. The portfolio's annualised returns and risk are then calculated for final assessment and comparison.

Algorithm 10 *Mean Variance*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$ where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
$\Pi_{(t)}$ is portfolio at current time, , where $\Pi_{(t)} = \langle \mathbf{c}_{(t)}, \mathbf{x}_{(t)} \rangle$.
optimiser() is an optimiser that uses quadratic programming to find the optimal portfolio.
mean() is a function that calculates the mean.
covar() is function that calculates the covariance.
reallocate() is a function that changes the proportion of capital invested in different QIMs according there weights, $\langle \Pi_t, \mathbf{w}_{(t)} \rangle$.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

$\mathbf{w}_{(0)} \leftarrow 0$
$\mathcal{P}_{(0)} \leftarrow 0$

for $t = 1$ to T **do**
  for $i = 1$ to n **do**
  $\mathbf{x}_{(1 \dots t,i)} \leftarrow$ get prices
  $\mathbf{m}_{(i)} \leftarrow mean(\mathbf{x}_{(1 \dots t,i)})$
  $\mathbf{\Sigma}_{(t,i)} \leftarrow covar(\mathbf{x}_{(1 \dots t,i)}, \mathbf{x}_{(1 \dots t,i)})$
  $\mathbf{w}_{(t,i)} \leftarrow optimiser(\mathbf{m}_{(i)}, \mathbf{\Sigma}_{(t,i)})$
  $\mathbf{c}_{(t,i)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t,i)})$
  $\mathcal{P}_{(t)} \leftarrow \mathcal{P}_{(t)} + \mathbf{x}_{(t,i)} \times \mathbf{c}_{(t,i)}$
  end for
end for

6.2.1.3 Fractional Kelly

The Fractional Kelly (FK) is a method for building a portfolio of simultaneous investment opportunities for which Kelly fractions have been calculated. However in the case of finance, Kelly has been known to give large estimates, recommending the use of leverage (MacLean et al., 2011). Hence the Kelly fractions have been scaled down uniformly so that the weights do not exceed unity, as we do not use leverage in our meta models, even though Kelly suggests that we do so. An analysis of Fractional Kelly was presented by MacLean et al. (2011), and Rising and Wyner (2012). In this model we estimate

Kelly for the underlying QIMs and use the Kelly fraction as the weights for capital allocation which leads to the meta portfolio.

To construct a Fractional Kelly portfolio (see Algorithm 11), we use the Fractional Kelly Function (see Algorithm 5). At each time step we calculate the Fractional Kelly weights. Once we have the Fractional Kelly weights, we can then use these weights to allocate capital for the next time step $(t + 1)$, which gives us the weighted return and hence the portfolio return. Fractional Kelly will serve as our third benchmark.

Algorithm 11 *Fractional Kelly*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\Pi_{(t)}$ is portfolio at current time, where $\Pi_{(t)} = \langle \mathbf{c}_{(t)}, \mathbf{x}_{(t)} \rangle$.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
$KELLY()$ is function described in Algorithm 5 that calculate fractional Kelly weights for all the QIMs at t.
$reallocate()$ is a function that changes the proportion of capital invested in different QIMs according there weights, $\langle \Pi_{(t)}, \mathbf{w}_{(t)} \rangle$.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

$\mathbf{w}_{(0)} \leftarrow 0$
$\mathcal{P}_{(0)} \leftarrow 0$

for $t = 1$ to T **do**
  $\mathbf{w}_{(t)} \leftarrow KELLY(\mathbf{x}_{(t)})$
  $\mathbf{c}_{(t)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t)})$
  $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \mathbf{x}_{(t)} \mathbf{c}_{(t)}$
end for

6.2.1.4 Optimal Kelly

In the case of Optimal Kelly (OK) set out in Algorithm 12, we maximise the log of wealth, using the quasi-newton approach with Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. Since the objective of Kelly is to maximise the logarithm of wealth, we extract optimal weights that maximise log of wealth using Equation 6.3,

$$\max_{\vec{w}} \sum_{t=1}^{T} \log\left(\sum_{i=1}^{n} 1 + (w_{i(t)} r_{i(t)}) \right). \tag{6.3}$$

Here r_i represents the returns of our QIMs and w_i represents the weights that will maximise the log of wealth. We use these weights to compute the return of the portfolio at the next time step $(t+1)$. Optimal Kelly will serve as our fourth benchmark.

Algorithm 12 *Optimal Kelly*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
$\Pi_{(t)}$ is portfolio at current time, where $\Pi_{(t)} = \langle \mathbf{c}_{(t)} \rangle$
reallocate() is a function that changes the proportion of capital invested in different QIMs according there weights, $\langle \Pi_{(t)}, \mathbf{w}_{(t)} \rangle$.
optimiser the optimiser uses the quasi-newton approach with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots . & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

$\mathbf{w}_{(0)} \leftarrow 0$
$\mathcal{P}_{(0)} \leftarrow 0$

for $t = 1$ to T **do**
  $\mathbf{w}_{(t)} \leftarrow optimiser(\mathbf{x}_{(t)})$
  $\mathbf{c}_{(t)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t)})$
  $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \mathbf{x}_{(t)} \mathbf{c}_{(t)}$
end for

6.2.2 New Meta Models

We will now present the meta models that we have developed. We have five models and we will present them in detail and explain why we choose a certain approach and how it is going to help us achieve our goals of generating better risk adjusted returns.

6.2.2.1 Median Kelly using Median of Returns

In Median Kelly (MK) we use the median of the data. The median of a distribution, just as in the case of the mean, is another way of assessing the central tendency of a distribution. However when data is not normally distributed the mean of the distribution may not necessarily be the best estimate of the central tendency. Previous research done by us on understanding forecast earnings had shown that median can be a better estimate in terms of giving better performance. Kelly, on the other hand is a distribution free approach and it is myopic in nature (see 2.2.4). Our objective here is to see if the median can give is a better Sharpe ratio and hopefully capture the central tendency better.

The model operates exactly like the Fractional Kelly model (see Section 6.2.1.3) except that the Kelly calculations are done using the median of the data signified by $\tilde{x}$ (Equation 6.4) to estimate Kelly instead of the mean as previously shown in Equation 2.9, all other parameters being the same. We then estimate the allocation percentage as we did for Fractional Kelly, but using the Median Kelly Function as shown in Algorithm 6. The return of the portfolio is then calculated by calculating the weighted return of the QIMs (see Algorithm 13),

$$f^* = \frac{\tilde{x} - r}{\sigma^2}. \tag{6.4}$$

Algorithm 13 *Median Kelly*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
$\Pi_{(t)}$ is portfolio at current time, where $\Pi_{(t)} = \langle \mathbf{c}_{(t)}, \mathbf{x}_{(t)} \rangle$.
$MEDIAN\ KELLY()$ is function described in Algorithm 6, it calculates Median Kelly weights for all the QIMs at t.
$reallocate()$ is a function that changes the proportion of capital invested in different QIMs according to their weights, $\langle \Pi_{(t)}, \mathbf{w}_{(t)} \rangle$.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

$\mathbf{w}_{(0)} \leftarrow 0$
$\mathcal{P}_{(0)} \leftarrow 0$

for $t = 1$ to T **do**
  $\mathbf{w}_{(t)} \leftarrow MEDIAN\ KELLY(\mathbf{x}_{(t)})$
  $\mathbf{c}_{(t)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t)})$
  $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \mathbf{x}_{(t)} \mathbf{c}_{(t)}$
end for

6.2.2.2 Kelly with Kalman Filter

One of the key challenges we face is changing data regime. A change in data regime can render a QIM loss-making, if the pattern or structure that it is targeting pauses for some time, potentially to restart again. A prudent portfolio manager or asset allocator will want to avoid a loss-making situation and will also prefer a situation where the capital is allocated efficiently whilst avoiding losses. Avoiding loss-making investments could improve both returns and Sharpe ratio. By avoiding loss making investments in a certain QIM and focusing on investments with potentially positive return, we are making use of the myopic property of Kelly; i.e. we only need to focus on the current best opportunity (see Section 2.2.4).

To overcome this challenge we make use of a Kalman Filter (see Section 2.5). We use the Kalman Filter to assess whether at $t+1$ our models are expected to return a positive or negative return. Specifically, we are not interested in the accuracy of return itself but in the sign of the return. We want our Kalman Filter to help us avoid periods of negative forecast returns but invest as usual using Fractional Kelly when the forecast is for a positive return (see Section 6.2.1.3).

To build this model we use three previously built functions, namely Fractional Kelly function, the Kalman Filter function and the performance curve function, presented in Algorithms 5, 8 and 7. To estimate the Kalman Filter for this model, we take the returns of each of the QIMs and generate a performance curve with base of 100, using Algorithm 7. Just as in the momentum model, our Kalman Filter is a simple scalar model i.e. there is just one variable: the performance curve of the QIM. Our approach is that the current observation value (at time t) is the best estimate of the future value (at time $t + 1$), with some error. Hence our state model coefficient is set to unity (1) and the variance is estimated by fitting a polynomial to the first 12 observation (one calendar year) in-sample period. The variance is then set to 1/10 of its value.

Once we have the Kalman Filter set up, we operate our model as shown in Algorithm 14 called Kelly with Kalman Filter (KKF). At each time step the Kalman Filter function forecasts which QIM will have a positive or negative return. QIMs with a positive forecast are assigned 1 and with a negative negatives 0. Calculate from Fractional Kelly weights using Fractional Kelly function. Remove any QIM with a negative forecast. Renormalising the Kelly weights recalculates the optimal Fractional Kelly weights, given the number of QIMs to invest in. The return of the portfolio is then determined by calculating the weighted return of the QIMs.

6.2.2.3 Median Kelly with Kalman Filter

The Median Kelly model with Kalman Filter (MK KF) is set up in the same way as the one presented above using Kelly (see Section 6.2.2.2), with the key difference being that Kelly estimates are calculated using the median of the distribution through the Median Kelly Function rather than the mean. Hence for this model we use the Median Kelly function seen earlier (Section 6.1.2). This model operates exactly the same way as represented in Algorithm 14. The steps are detailed in Algorithm 15. The Key difference between Algorithm 15 and 14 is that we use the Median Kelly Function (Algorithm 6) instead of the Fractional Kelly Function (Algorithm 5).

Algorithm 14 *Kelly with Kalman Filter*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
$\Pi_{(t)}$ is portfolio at current time, where $\Pi_{(t)} = \langle \mathbf{c}_{(t)}, \mathbf{x}_{(t)} \rangle$
reallocate() is a function that changes the proportion of capital invested in different QIMs according to their weights, $\langle \Pi_{(t)}, \mathbf{w}_{(t)} \rangle$.
PERFORMANCE CURVE() is the function described in Algorithm 7. It converts QIM returns to price.
KELLY() is a function described in Algorithm 5 it calculates fractional Kelly weights.
BINARY KALMAN FILTER() is a function described in Algorithm 8. It gives a binary output based on the forecast from our Kalman Filter.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

for $t = 1$ to T **do**
  $Kelly_{(t)} = KELLY(\mathbf{x}_{(t)})$
  $sum \leftarrow 0$
  for i = 1 to n **do**
    $price_{(t,i)} \leftarrow PERFORMANCE\ CURVE(\mathbf{X}_{(t,i)})$
    if $BINARY\ KALMAN\ FILTER(price_{(t,i)})! = 1$ **then**
      $Kelly_{(t,i)} = 0$
    end if
    $sum \mathrel{+}= Kelly_{(t,i)}$
    for $i = 1$ to n **do**
      $Kelly_{(t,i)} = \frac{Kelly_{(t,i)}}{sum}$
    end for
  end for
  $\mathbf{w}_{(t)} \leftarrow Kelly_{(t)}$
  $\mathbf{c}_{(t)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t)})$
  $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \mathbf{x}_{(t,i)} \mathbf{c}_{(t,i)}$
end for

Algorithm 15 *Median Kelly with Kalman Filter*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
$\Pi_{(t)}$ is portfolio at current time, where $\Pi_{(t)} = \langle \mathbf{c}_{(t)}, \mathbf{x}_{(t)} \rangle$.
reallocate() is a function that changes the proportion of capital invested in different QIMs according to their weights, $\langle \Pi_{(t)}, \mathbf{w}_{(t)} \rangle$.
PERFORMANCE CURVE() is a function described in Algorithm 7, it converts QIM returns to price.
MEDIAN KELLY() is a function described in Algorithm 6. It calculates median Kelly weights.
BINARY KALMAN FILTER() is a function described in Algorithm 8. It gives us a binary output based on the forecast from our Kalman Filter.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

for $t = 1$ to T **do**
  $Kelly_{(t)} = MEDIAN\ KELLY(\mathbf{x}_{(t)})$
  $sum \leftarrow 0$
  for i = 1 to n **do**
    $price_{(t,i)} \leftarrow PERFORMANCE\ CURVE(\mathbf{X}_{(t,i)})$
    if $BINARY\ KALMAN\ FILTER(price_{(t,i)})! = 1$ **then**
      $Kelly_{(t,i)} = 0$
    end if
    $sum \mathrel{+}= Kelly_{(t,i)}$
    for $i = 1$ to n **do**
      $Kelly_{(t,i)} = \frac{Kelly_{(t,i)}}{sum}$
    end for
  end for
  $\mathbf{w}_{(t)} \leftarrow Kelly_{(t)}$
  $\mathbf{c}_{(t)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t)})$
  $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \mathbf{x}_{(t,i)} \mathbf{c}_{(t,i)}$
end for

6.2.2.4 Fractional Kelly with Simple Moving Average

The moving average has been very robust in many complex situations; it has been most extensively used to track targets and to smooth data serving as a low pass filter. The moving average is also very adaptable with no distributional assumptions and different ways to calculate it, hence giving room for customization as and when the need arises. In this model, as in the case of the two previous models (see Section 6.2.2.2 and 6.2.2.3) we are aiming to avoid investing in QIMs that are potentially loss-making and in the process making optimal allocation of capital. In this process we aim to improve our returns and our Sharpe ratio. Once again, we are making use of the myopic property of Kelly; i.e. we only need to focus on the current best opportunity (see 2.2.4).

In this model we use Fractional Kelly function and the performance curve function, shown in Algorithms 5 and 7. We elaborate the steps set out in Algorithm 16 called Fractional Kelly with Moving Average (K MA). We initialise and create the performance curve of each QIM that we have and calculate Fractional Kelly for each QIM. We then create a simple moving average using the performance curve for each QIM. For each model we check if the performance curve is equal or higher than the moving average. If so we want to allocate capital to that QIM and the signal is 1. If it is below the moving average then the signal is 0 indicating that we don't want to invest in that QIM as it is likely to be loss-making. Specifically a QIM that is above the moving average indicates continued positive performance and vice versa

For QIMs whose signal is 0, the Kelly weight is also set to 0. Subsequently, we renormalise the weights giving us the optimal Fractional Kelly weights for QIMs with positive expected return. The renormalised weights are now used to allocated capital for the next period $(t + 1)$ and calculate portfolio returns.

6.2.2.5 Median Kelly with Simple Moving Average

The Median Kelly with Moving Average algorithm is very similar to the Kelly with Moving Average algorithm presented above (Section 6.2.2.4), with the key difference being that in this model we use median of the data to estimate Kelly through the Median Kelly Function 6.1.2 (line 4). Besides that, this algorithm works exactly the same as Algorithm 16 discussed above. This algorithm is called Median Kelly with Simple Moving Average (MK MA) and is elaborated in Algorithm 17.

Algorithm 16 *Kelly with Simple Moving Average*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
$\Pi_{(t)}$ is portfolio at current time. where $\Pi_{(t)} = \langle \mathbf{c}_{(t)}, \mathbf{x}_{(t)} \rangle$
$reallocate()$ is a function that changes the proportion of capital invested in different QIMs according to their weights, $\langle \Pi_{(t)}, \mathbf{w}_{(t)} \rangle$.
$moving\_average()$ is a function that calculates the moving average for the QIMs using the performance curve of the QIMs.
$PERFORMANCE\ CURVE()$ is a function described in Algorithm 7. It converts QIM returns to price.
$KELLY()$ is a function described in Algorithm 5. It calculates median Kelly weights.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

for $t = 1$ to T **do**
    $Kelly_{(t)} = KELLY(\mathbf{x}_{(t)})$
    $sum \leftarrow 0$
    for i = 1 to n **do**
        $x_{(t,i)} \leftarrow$ get current price of QIM.
        $p_{(t,i)} \leftarrow PERFORMANCE\ CURVE(\mathbf{X}_{(t,i)})$
        $y_{(t,i)} \leftarrow moving\_average((p_{(1\dots t,i)}))$
    if $p_{(t,i)} \geq y_{(t,i)}$ **then**
        $Kelly_{(t,i)} = 1$
    else
        $Kelly_{(t,i)} = 0$
    end if
    $sum \mathrel{+}= Kelly_{(t,i)}$
    for $i = 1$ to n **do**
        $Kelly_{(t,i)} = \frac{Kelly_{(t,i)}}{sum}$
    end for
    end for
    $\mathbf{w}_{(t)} \leftarrow Kelly_{(t)}$
    $\mathbf{c}_{(t)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t)})$
    $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \mathbf{x}_{(t,i)} \mathbf{c}_{(t,i)}$
end for

Algorithm 17 *Median Kelly with Simple Moving Average*: *Input* = QIM returns, *Output* = Portfolio returns.

READ QIM returns as $\mathbf{X}_{(n,t)}$, where t is time stamp and n is the number of QIM models.
$\mathcal{P}_{(t)}$ is the value of the of the portfolio at t.
$\mathbf{w}_{(t)}$ are the weights to allocate capital to the QIMs.
$\mathbf{c}_{(t)}$ investment capital.
Π_t is portfolio at current time, where $\Pi_{(t)} = \langle \mathbf{c}_{(t)}, \mathbf{x}_{(t)} \rangle$
$PERFORMANCE\ CURVE()$ is a function described in Algorithm 7. It converts QIM returns to price.
$reallocate()$ is a function that changes the proportion of capital invested in different QIMs according to their weights, $\langle \Pi_{(t)}, \mathbf{w}_{(t)} \rangle$.
$moving\_average()$ is a function that calculate the moving average for the QIMs using the performance curve of the QIMs.
$MEDIAN\ KELLY()$ is a function described in Algorithm 6 that calculates median Kelly weights.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{tn} \\ x_{21} & x_{22} & \dots & x_{tn} \\ \dots & & & \\ x_{T1} & x_{T2} & \dots & x_{Tn} \end{bmatrix}$$

for $t = 1$ to T **do**
  $Kelly_{(t)} = MEDIAN\ KELLY(\mathbf{x}_t)$
  $sum \leftarrow 0$
  for i = 1 to n **do**
    $x_{(t,i)} \leftarrow$ get current price of QIM.
    $p_{(t,i)} \leftarrow PERFORMANCE\ CURVE(\mathbf{X}_t)$
    $y_{(t,i)} \leftarrow moving\_average((p_{(1 \dots t,i)}))$

    if $p_{(t,i)} \geq y_{(t,i)}$ **then**
      $Kelly_{(t,i)} = 1$
    else
      $Kelly_{(t,i)} = 0$
    end if
    $sum \mathrel{+}= Kelly_{(t,i)}$
    for $i = 1$ to n **do**
      $Kelly_{(t,i)} = \frac{Kelly_{(t,i)}}{sum}$
    end for
  end for
  $\mathbf{w}_{(t)} \leftarrow Kelly_{(t)}$
  $\mathbf{c}_{(t)} \leftarrow reallocate(\Pi_{(t-1)}, \mathbf{w}_{(t)})$
  $\mathcal{P}_{(t)} \leftarrow \sum_{i=1}^{n} \mathbf{x}_{(t,i)} \mathbf{c}_{(t,i)}$
end for

6.3 Summary

In this chapter we presented the five meta models based on Kelly and four benchmarks that allocate capital to the QIMs:

- Equally weighted,
- Mean Variance Optimisation,
- Fractional Kelly model,
- Optimal Kelly model,
- Median Kelly,
- Fractional Kelly with Kalman Filter,
- Median Kelly with Kalman Filter,
- Fractional Kelly with Simple Moving Average, and
- Median Kelly with Simple Moving Average.

We chose Kelly because it gives us the optimal investment size for an investment with positive expectation. At the same time, we were cognizant of some very real challenges that an asset allocator faces when dealing with heterogeneous data; hence risk controls are important. We show how we build several models with the objective of not only getting better Sharpe ratio but also to address a very real problem faced in finance; that of heterogeneous data that can change any time and render a QIM loss making. We presented five models to show the evolution in our thinking to get to the best performing model, which we will see in Chapter 7 and Chapter 8 when we test them with synthetic and real financial data. The meta models presented in this chapter address two important requirements, correct investment size and risk control which, as we discussed in Chapter 1 are important components of good portfolio management. In the next chapter we will generate some synthetic data to test our models and see if we are on the correct path.

Chapter 7

Generating Synthetic Data for Models

In the previous two chapters, we presented a set of quantitative models that would interact with the markets (Chapter 5) and another set of models that do capital allocation to these models (Chapter 6), leading to our meta portfolio. Both set of models are key components of our two-tier framework. In this Chapter we will generate some synthetic data to see if these models that we have proposed in the last two chapters actually work. More importantly we also want to check if the two-tier framework that we have created actually gives us the benefits that we imagine it should.

We begin this chapter by generating some synthetic data that resembles prices of stocks, by using a GARCH model. We will ensure that the data has key features found in financial price data. For our synthetic data to resemble financial data we will ensure that our data has correlation, trends and jumps, all characteristics found in real stock prices. We will then run statistical tests on this data to identify these features and patterns in the synthetic data (see Chapter 3).

7.1 Synthetic Data to build Quantitative Investment Models

The purpose of this synthetic data is to check whether conceptually we are on the right track. Specifically we want to see whether our QIMs will capture momentum and mean reversion as we expect them to (see Section 2.3). Furthermore we also want to see if the meta models will perform as we expect; i.e. will we get better risk adjusted returns as we expect them to.

To achieve our objective we will generate data that actually resemble financials returns and induce certain characteristics which would make it fairly straightforward for our

QIMs to capture these patterns. For example, trend in the data will help the momentum model and higher variance and correlation would help the MN Pairs and long only models. To achieve this use a Generalised Autoregressive Conditional Heterosekdasticity (GARCH) model, and then we will induce trends into the data. We first present the GARCH model.

A GARCH model has three components shown in equation Equation 7.1,

$$\sigma_n^2 = \gamma V_l + \sum_{i=1}^{q} \alpha u_{n-i}^2 + \sum_{i=1}^{p} \beta \sigma_{n-i}^2. \tag{7.1}$$

Specifically the the three components are, i) V_l represents long-term variance, ii) u^2 represents squared error terms over a specified window length and iii) α is lagged variance again over a specified window length. Their respective coefficients γ, α and β represent their weight. GARCH(p, q) model is specified as such, p is the number of lags for the conditional variance and q is the number of lags for the squared error term. The most widely used GARCH models are GARCH(1,1) models, here the value of p and q is 1; we will also use these models. A typical GARCH model is shown in Equation 7.1 and a specimen output in Figure 7.2, we can also see a real world specimen of GARCH Figure 7.3 for the NASDAQ 100 Index.

Our objective is to generate a data set for our QIMs so that we can construct portfolios. For this purpose we need to create data sets that have correlations found in stocks that belong to the same sector. Hence we need a multivariate version of the model presented in Equation 7.1 for our simulation. To this end we will use the Constant Conditional Correlation Multivariate (CCC-MVGARCH). This model was proposed by Bollerslev (1990). It has one extra feature, it imposes correlation on the data set.

A multivariate GARCH model can be defined as $\mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{a}_t$ and $\mathbf{a}_t = \mathbf{H}_t^{\frac{1}{2}}\mathbf{z}_t$. Here $\boldsymbol{r}_t$ = n x 1 vector of returns at time t, $\boldsymbol{a}_t$ = n x 1 of mean corrected returns of n assets at time t, i.e. $E[\mathbf{a}_t] = 0$, $\text{Cov}[\mathbf{a}_t] = [\mathbf{H}_t]$. $\boldsymbol{\mu}_t = n$ x 1 vector of expected value of the conditional r_t. $\boldsymbol{H}_t = n$ x n matrix of conditional variance of $\boldsymbol{a}_t$ at time t, $\mathbf{H}_t^{\frac{1}{2}} = n$ x n matrix at time t such that $\mathbf{H}_t$ is the conditional variance matrix of a_t. $\mathbf{H}_t^{\frac{1}{2}}$ can be obtained by Cholesky factorisation of $\mathbf{H}_t$ and $\mathbf{z}_t = n$ x 1 of i.i.d errors such that $E[\mathbf{z}_t]$ = 0 and $E[\mathbf{z}_t\mathbf{z'}_t] = I$.

A CCC-MVGARCH model is built using conditional variance as well as correlations. The conditional covariance matrix $\boldsymbol{H}_t$ is made as such:

$$\boldsymbol{H}_t = \boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t \tag{7.2}$$

Where $\boldsymbol{D}_t = diag(\boldsymbol{h}_{1t}^{\frac{1}{2}}, \ldots, \boldsymbol{h}_{nt}^{\frac{1}{2}})$ is the conditional standard deviation and $\boldsymbol{R}_t$ is the conditional correlation matrix, hence becomes $\boldsymbol{R}$ and Equation 7.2 becomes:

$$\boldsymbol{H}_t = \boldsymbol{D}_t \boldsymbol{R} \boldsymbol{D}_t \tag{7.3}$$

The correlation matrix $\boldsymbol{R} = [\rho_{ij}]$ is positive definite with $[\rho_{ii}] = 1,\ i = 1, \ldots, n$. The off-diagonals elements of the conditional covariance matrix $\mathbf{H}_t$, are given by:

$$[\mathbf{H}_t]_{ij} = \rho_{ij}\sqrt{h_{it}h_{jt}} \qquad i \neq j \tag{7.4}$$

The process a_{it} is modelled as univariate GARCH. Hence the conditional variance can be written in a vector form:

$$\mathbf{h}_t = \mathbf{c} + \sum_{j=1}^{q} \mathbf{A}_j \mathbf{a}_{t-j}^2 + \sum_{j=1}^{p} \mathbf{B}_j \mathbf{h}_{t-j}, \tag{7.5}$$

Where $\mathbf{c}$ is n x 1 vector, $\mathbf{A}_j$ and $\mathbf{B}_j$ are diagonal n x n matrices, and $\mathbf{a}_{t-j}^2 = \mathbf{a}_{t-j} \odot \mathbf{a}_{t-j}$ is the element-wise product. $\mathbf{H}_t$ is ensured positive definite when the elements of $\mathbf{c}$ and $\mathbf{A}_j$ and $\mathbf{B}_j$ are positive and $\mathbf{R}$ is positive definite.

Parameters for Simulation

For our synthetic data we have the following parameters, as shown in the univariate case (see Equation 7.1 to Equation 7.5) $\alpha = 0.70$ representing volatility from the previous period, $\beta = 0.20$ representing variance from the previous time step and the intercept $\gamma = 0.20$ representing constant minimum volatility, this set in $\mathbf{D}$. We choose these parameters as they gave us enough variance for the MN pairs model to generate enough trades. The correlation is 0.75 set in $\mathbf{R}$ to give correlated data to resemble sector pairs. Both p and q are set to 1 as we have GARCH(1, 1) model.

We generate 100 data sets; each data set has 12 correlated time series from the CCC-GARCH model. We choose to go with 12 time series as we will have enough data to make pairs for the MN pairs models as well as sector groupings for the long only model. Each of the 12 time series will have 1000 observations. This is approximately four years in time since each year has approximately 252 working days. Variance for each of the 12 time series is [1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1] and the mean is set at 0.1 so that the prices don't go negative. The variance increases linearly as it needs to be different enough for the models to differentiate between data points, but still group them as similar, as in the case of sectors.

Our synthetic data now has GARCH characteristics correlation of 0.75 and now our next step is induce trends in the data. We do so by skewing the mean. We add 0.35 to each of the first 200 data points and then −0.25 to the next 200, alternating to the end of the data set. This shift in mean ensures that will have trends in the data instead of just Brownian motion. Here we chose 200-day window so that we can get a trend going long enough to make our QIM profitable. A specimen of synthetic prices can be seen in Figure 7.1.

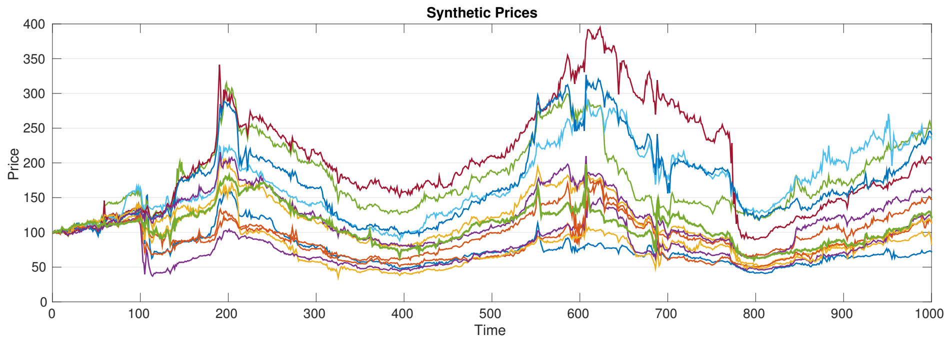


FIGURE 7.1: Prices generated from using synthetic GARCH data and skewed mean giving the data trends after 200 data points. This figure shows the batch of 12 price time series number 25 to 36.

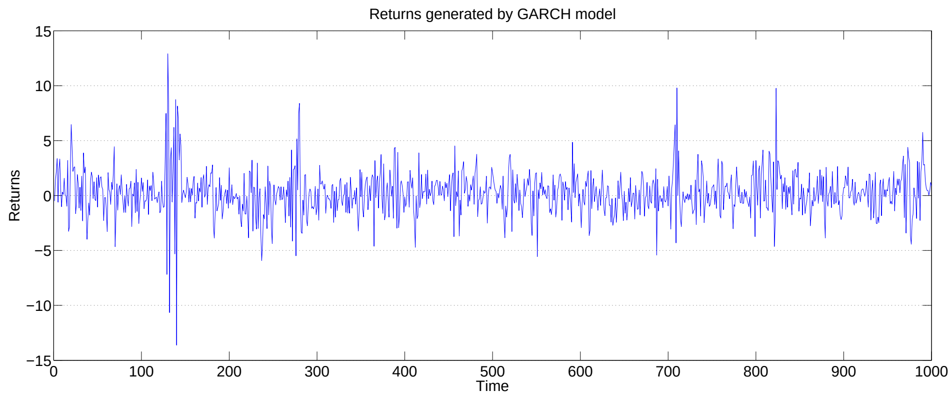


FIGURE 7.2: Synthetic GARCH data, showing volatility clustering, high volatility periods followed by high volatility and low volatility periods by low volatility. This figure depicts the GARCH effect, used for modelling. This data series is the time series numbered 6 in the first block of GARCH data.

Once the data has been generated we convert them into stock prices by giving them a base price of 100. These artificial time series are now ready to be used for statistical testing and building our four quantitative investment models.

Parameters of the QIMs

Now that we have the synthetic data ready we are ready to structure the data to be used by the QIMs as set out in Chapter 5. The QIMs will be set up just as in the case of

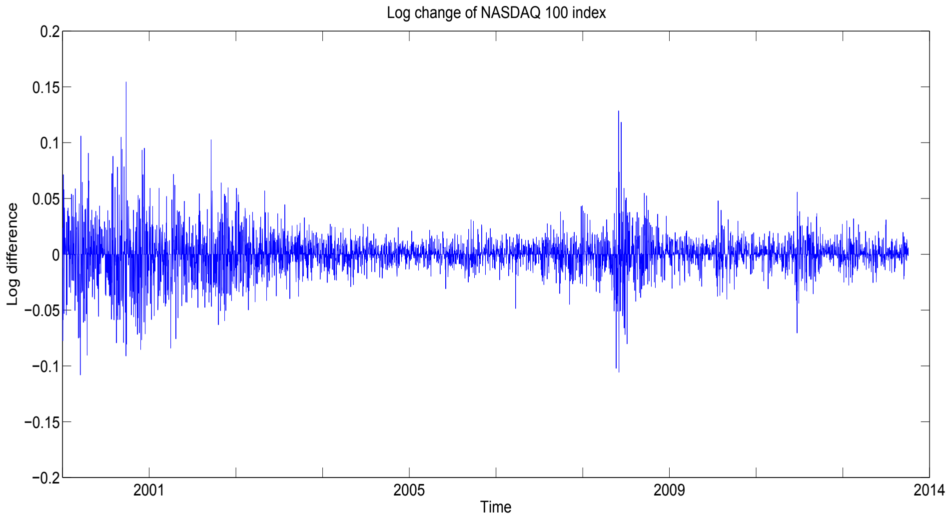


FIGURE 7.3: NASDAQ 100 index daily price change from year 2000 to the middle of 2013, showing volatility clustering, high volatility periods followed by high volatility and low volatility periods by low volatility depicting the GARCH effect.

real financial data, i.e. we will create sector groups and pairs as required by the relevant QIM.

Momentum Model Parameters

The momentum model has one key parameter the window of the moving average, which we set at 100 observations approximately five months, and half the length of the 200 period skew we impose on the data. We chose a 100 period window after trying different window sizes. The 100 period window gave us strong positive returns and a high Sharpe ratio. As we will see in the next chapter with real financial data also we use a 100 period window.

MN Pairs Model Parameters

The MN pairs model, as we know, operates on paired data. Hence for this model we need to make pairings from the data we have from the simulation. Our pairing need to have similar variance so as to mimic the behaviour of stocks from the same sector. Since in each batch we have 12 time series we will make 11 pairs as such. Let n the be first time series and $n+11$ the last time series of the the set of 12. We make pairs from the closest times series i.e. the subsequent time series, $(n \to n+1)$, $(n+1 \to n+2)$ ... $(n+10 \to n+11)$. Given that the variance increases linearly the adjacent time series will have similar behaviour. This model operates on a 10 period window. The speed of

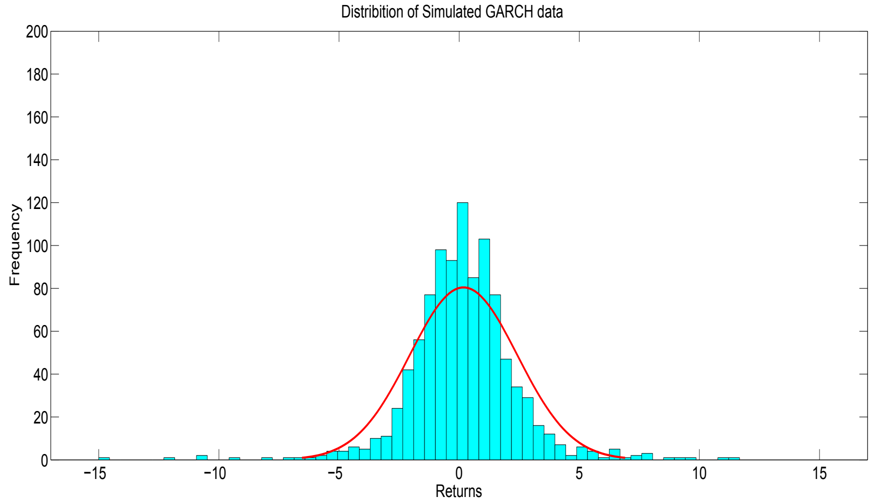


FIGURE 7.4: In blue we can see the distribution of the synthetic GARCH data, compared to the theoretical normal distribution curve in red, showing that the data is not normally distributed and exhibits skew and kurtosis. It fails the Jarque-Bera test at 5% significance level with the test statistic = 1176.58408. Mean = 0.181699, variance = 2.242049, skew = −0.198431 and kurtosis = 9.393573. Bin size is calculated using Freedman - Diaconis rule and has 59 bins, bin size is 0.452203. This data is the first time series of the first run of GARCH simulation.

mean reversion as measured by λ in the OU model is Γ set at 0.1, and standard deviation threshold δ is set at 2. This gives us the maximum number of profitable trades.

Long Only Model

The long only model operates on the basis of sector groupings. Since in each batch we have 12 time series we will create three sector groups each containing four time series each. Let n the be first time series and $n+11$ the last time series of the the set of 12. We make groups from the closest times series i.e. the closest four time series, $(n \to n+3)$, $(n+4 \to n+8)$ $\ldots$ $(n+9 \to n+12)$.

7.2 Performance of Models

The output of the QIMs is used by the meta models to form the meta portfolio, which is a portfolio of QIMs. The meta models are our main contribution in this thesis. The objective of our meta models is to maximise our Sharpe ratio. We will now present the performance of these models i.e. the three QIMs and nine meta models, including their four benchmarks.

We begin with the results of the QIMs are followed by the results of the meta models. In the case of QIMs we find that on average the models are profitable as shown in Figure 7.6

and Table 7.1, where we can see the returns, volatility and Sharpe ratios. Out of 100 runs the momentum and market-neutral pairs models generate positve returns on each run, whereas the long only model has 10 losing runs. The momentum model is the best performer, with the highest returns, as expected since it has the the best opportunity to thrive with 12 time series in each simulation run. Both momentum and long only have the highest volatility, as they both make investments that are not offset on the other side like the MN pairs model.

The MN pairs model has the lowest volatility, as expected, since every investment buys and sells a stock at each trade, neutralising much of the volatility. The low volatility of the MN pairs model helps achieve the highest Sharpe ratio as well and this is natural for an approach such as MN Pairs (see Figure 7.7). In Figure 7.5 we can see the distribution of the Sharpe ratio for all the the QIMs after 100 runs. They are reasonably well behaved; the long-only model is the only model that generates some some negative returns hence negative Sharpe ratio. Overall, the results are strong and at the same time not a surprise since we engineered the data to have characteristics that would aid the QIMs.

Table 7.1: Performance of the quantitative investment models

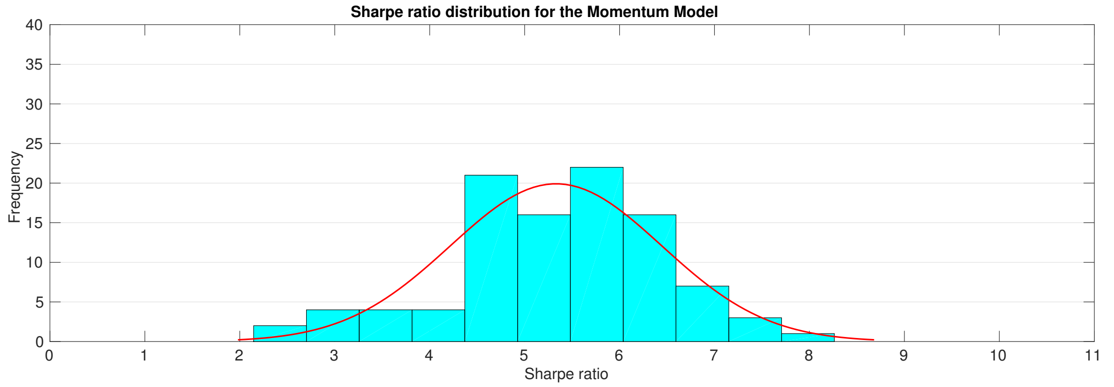
	Momentum	Long Only	MN Pairs
Average Ann. return	145.59%	32.17%	33.31%
Average Ann. volatility	27.79%	37.78%	3.71%
Average Sharpe ratio	5.33	0.90	9.14

This table represents the average annualised return, Sharpe ratio and volatility of the QIMs over 100 simulation runs.

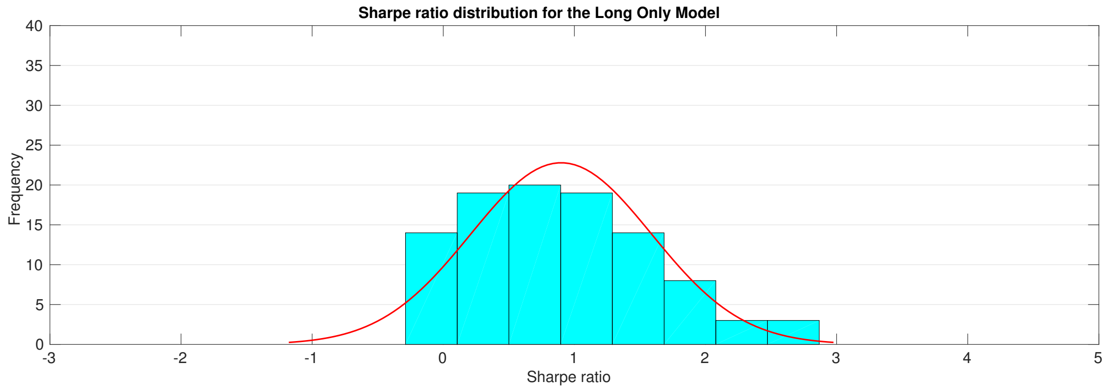
We now move to the meta models, which allocate capital to the QIMs and construct a meta portfolio. Among the meta models we have five models made by us and four benchmarks. We can see the Sharpe ratios of the models in Figure 7.8 and Table 7.3. Among the benchmarks, we expected MVO to have the best Sharpe ratio, but Fractional Kelly is the best performer and Optimal Kelly is the worst performer. From the five models that we have built, Median Kelly is the worst performer with the lowest Sharpe ratio of the five models. However the rest of the models beat the best benchmark model's Sharpe ratio by some margin.

Furthermore, when we analyse the Sharpe ratios on each of the runs we find that Median Kelly with Moving Average, Kelly with Moving Average and Kelly with Kalman Filter beat the benchmarks with high probability, as shown in Table 7.2. This is very promising not only for our model, but also for the framework.

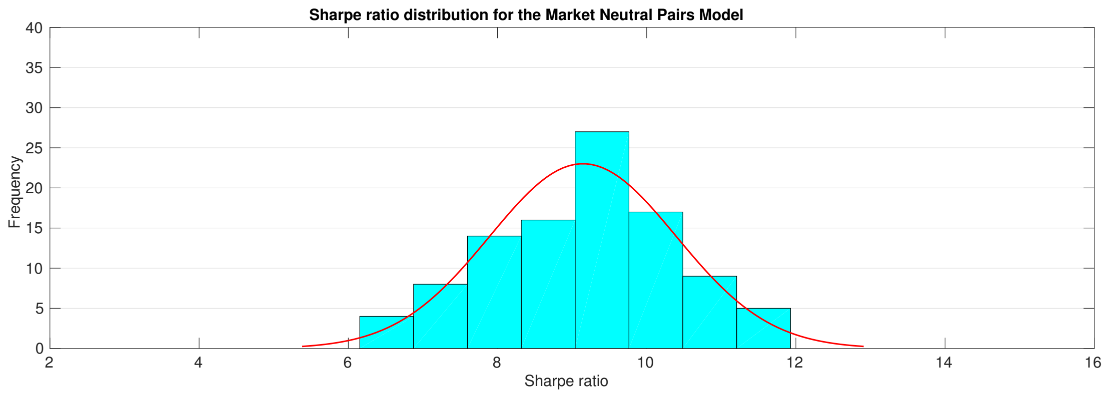
We assess the returns generated by our meta models in Figure 7.9 and Table 7.4. Among the benchmarks, the Optimal Kelly model has the best returns, beating all the other benchmarks. Equally Weighted model performs better than MVO and Fractional Kelly; this is owing to the strong performance shown by the QIMs. Furthermore MVO is the



(a) Momentum model's distribution of Sharpe ratio. Number of bins = 11, bin size = 0.5561, mean = 5.3310, variance = 1.2410, skew = −0.4060, kurtosis = 0.6565.



(b) Long Only model's distribution of Sharpe ratio.Number of bins = 8, bin size = 0.3944, mean = 0.8991, variance = 0.4767, skew = 0.5120, kurtosis = −0.1081.



(c) MN Pairs model's distribution of Sharpe ratio. Number of bins = 8, bin size = 0.7218, mean = 9.1451, variance = 1.5654, skew = −0.1379, kurtosis = −0.2738.

FIGURE 7.5: Distribution of Sharpe ratios of the QIMs after 100 runs. The plots are made using Freedman-Diaconis method.

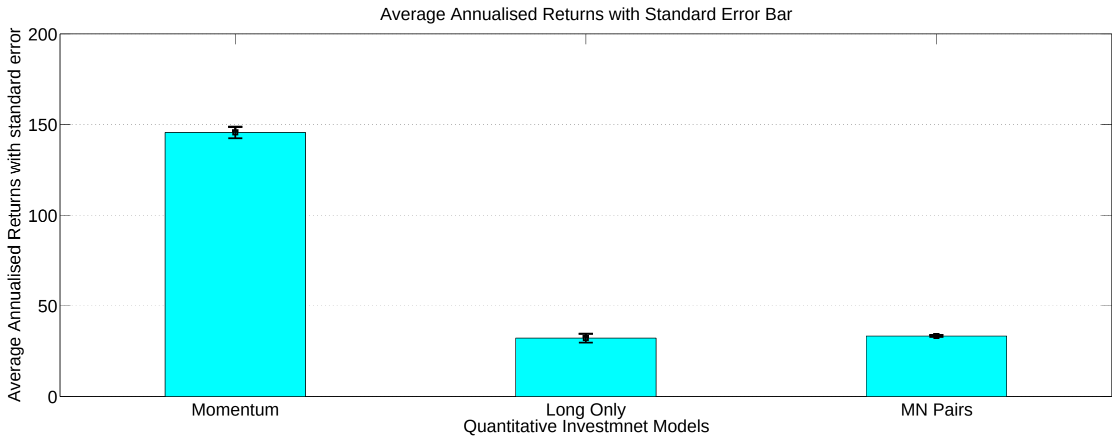


FIGURE 7.6: Annualised return with standard error of long only, momentum and MN pairs model using synthetic GARCH data with trend and correlation incorporated in the data. The average annualised return and standard error is calculated from 100 simulation runs.

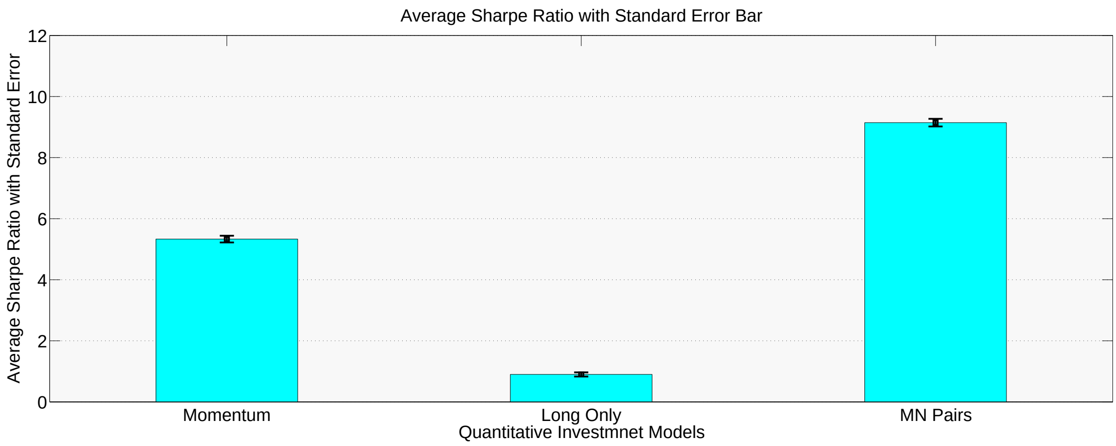


FIGURE 7.7: Sharpe ratio with standard error bars for the long only, momentum and MN Pairs model using synthetic GARCH data with trend and correlation incorporated in the data. The average Sharpe ratio and standard error is calculated from 100 simulation runs.

TABLE 7.2: Sharpe Ratio out-performance

	MK MA	K MA	MK MK	K KF	M K
MVO	89	93	84	91	37
F K	94	97	89	95	23
O K	99	99	96	97	92
Average	94	96	90	94	51

This table shows the number of times our meta models beats the key benchmarks in terms of Sharpe ratio with synthetic data, the Kelly models with moving average perform well.

worst performer. This is not surprising as it will always pick the portfolio with minimum variance. Fractional Kelly is in the middle of the pack. Among the five models that we have built, Median Kelly with Moving Average is the best performer, but it is still beaten by Optimal Kelly by a huge margin. This is primarily because the Optimal Kelly model tends to pick just a single best performer nearly always instead of building a portfolio. The performance of the rest of the models is pretty similar, except for the theme that Median Kelly models including the ones with Moving Average and Kalman Filter all perform better than the Kelly models. We did not expect this difference between median and mean Kelly performance.

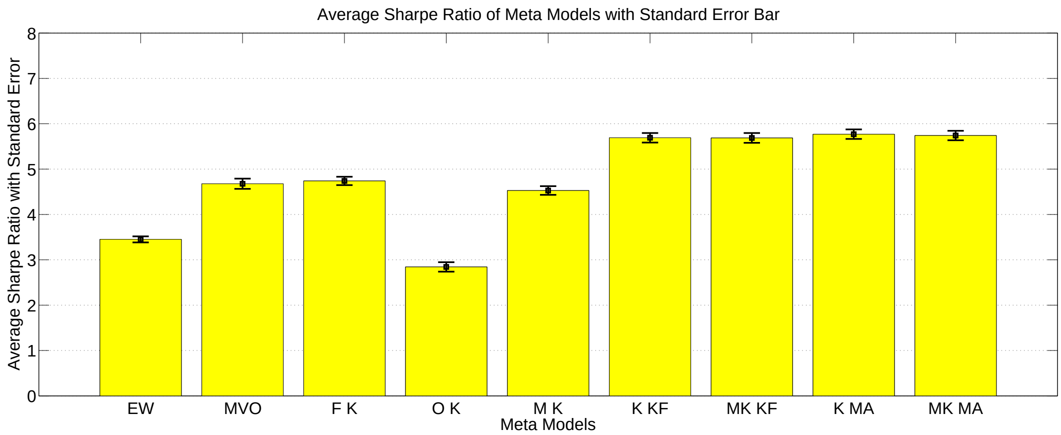


FIGURE 7.8: Sharpe ratio with standard error of four benchmarks Equally Weighted, Mean Variance, Fractional Kelly and Optimal Kelly, and five models we developed, Median Kelly, Kelly with Kalman Filter, Median Kelly with Kalman Filter, Kelly with Moving Average and Median Kelly with Moving Average using synthetic data with trend and correlation of 0.75 incorporated in the data. The average return and standard error is calculated from 100 runs of QIMs. Kelly with Moving Average and Median Kelly with Moving Average both beat our benchmarks, by a good margin.

TABLE 7.3: Sharpe Ratio

	EW	MVO	F K	OK	MK	K KF	MK KF	K MA	MK MA
Sharpe ratio	3.46	4.67	4.73	2.84	4.52	5.69	5.68	5.77	5.79

The table represents the average Sharpe ratio of the Meta models where the input is from QIM models that were using GARCH data set.

TABLE 7.4: Annualised returns

	EW	MVO	F K	OK	MK	K KF	MK KF	K MA	MK MA
Returns	57.65	49.94	52.12	91.24	54.59	52.98	55.09	53.63	56.32

The table represents the average annualised returns of the meta models where the input is from QIM models that were using GARCH data set.

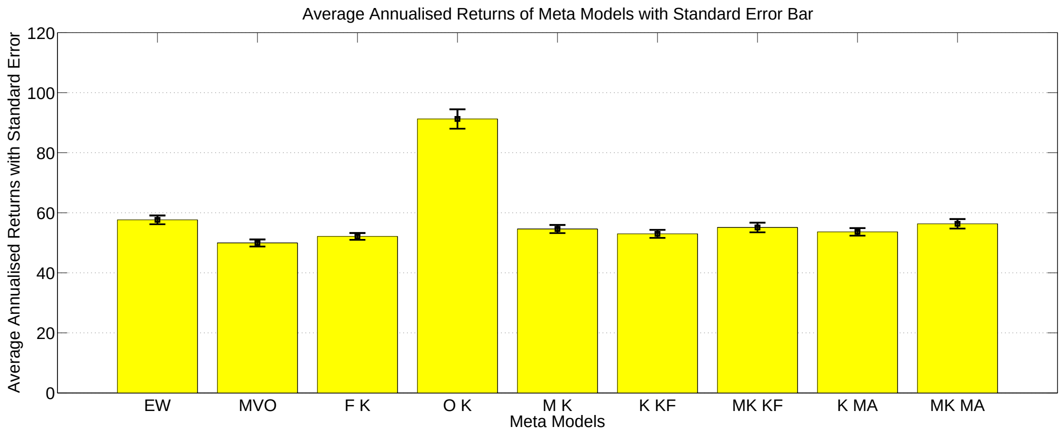


FIGURE 7.9: Annualised returns with standard error of four benchmarks Equally Weighted, Mean Variance, Fractional Kelly and Optimal Kelly, and five models we developed Median Kelly, Kelly with Kalman Filter, Median Kelly with Kalman Filter, Kelly with Moving Average and Median Kelly with Moving Average using synthetic data with trend and correlation of 0.75 incorporated in the data. The average return and standard error is calculated from 100 runs of QIMs. Optimal Kelly is the winner by large margin followed by Equally Weighted and Median Kelly with Moving Average.

7.3 Summary

In this chapter we generated some synthetic data with strong momentum and mean reversion features using a GARCH model that generated correlated data. We induced strong momentum in the data. The data was then used by our QIMs that are designed to build a portfolio capturing patterns that exist on differing time horizons, targeting momentum and mean reversion. The output of the QIMs is used by the meta models to construct a meta portfolio; this is done through capital allocation by the meta models. This was our first test to see each of the components of good portfolio management, namely signal, capital allocation and risk, being addressed together as discussed in Chapter 1.

This chapter has shown that it is possible to achieve a higher Sharpe ratio by constructing a meta portfolio of QIMs capturing different patterns in data at differing time horizons. We also saw in this chapter that there is merit to the two-tier framework that we introduced in Chapter 4. We now need to see if this is only possible with synthetic data or will this framework hold firm and give similar results with real financial data. In the next chapter well use real financial data to see if our models perform the way we have seen in the simulation.

Chapter 8

Performance Analysis of Models Using Real Data

In the previous chapter we presented our QIMs and meta models, using synthetic data generated using a GARCH model. Using synthetic data we found that Optimal Kelly gave us the best annualised returns and Median Kelly with Moving Average gave us the best Sharpe ratio. In this chapter we will test our meta models on real financial time series data to see how our models perform in this environment and whether we get similar comparative performance. First we shall present the data that we are going to use in Section 8.1. Then we discuss the steps we take to clean the data, ensuring our data points are consistent across all markets both for in-sample and out-of-sample results. In Section 8.2 we present the parameters of the the four QIM and how we use the data in these models.

Once the data is ready in Section 8.3 we present the in-sample performance of the QIMs, followed by the out-of-sample performance in Section 8.4. In Section 8.5 we present the in-sample and out-of-sample results of the meta models our main contribution. We follow this up with analysis of the results.

8.1 Markets

For our experiments we use real financial data from global stock exchanges. The source of our data is Bloomberg, which aggregates all financial data from global financial exchanges. The data starts from 2005 and ends in 2012, where data from 2005 to 2009 is treated as the in-sample data and 2010 to 2012 as the out-of-sample data. The in-sample data is used to build and parameterise the model, and the out-of-sample data is used for validation of the models. For equity indices we use global equity indices with active futures market (see Table 8.1). For individual stocks we have used the 250 biggest stocks in Europe, that are denominated in Euros, and we have added 100 of the largest stocks by capitalisation in UK, which are all FTSE 100 constituents. We present their names in the Appendix N.1.

Table 8.1: List of markets

Index	Country
EuroSTOXX	Europe
DAX	Germany
CAC40	France
MIB	Italy
IBEX	Spain
AEX	Netherlands
FTSE 100	UK
SMI	Switzerland
BEL 20	Belgium
BIST 30	Turkey
S&P 500	USA
NASDAQ 100	USA
S&P TSX 60	Canada
Bovespa	Brazil
Nikkei 225	Japan
Hang Seng	Hong Kong
S&P ASX 200	Australia

List of global markets that are the source of real stock price and equity index data, For equity index futures we use all of the above markets, for stocks we use UK, Germany, France, Italy, Spain, Netherlands and the broader European index Euro STOXX.

8.1.1 Data Cleaning

Financial data is large and can have errors and misprints. Since our strategies are heavily dependent on good quality data, it is imperative that our data is clean and aligned for our experiments. This is because statistical models are very sensitive to errors in data, and generally fail if they are not aligned properly.

- **Errors and misprints:** In large data sets some data can get corrupted owing to change in company name and the old data sets sometimes are not fully integrated into new ones. Sometimes there can be misprints in data. These errors need to be removed or corrected.

- **Gaps and alignment:** Where there are missing data points owing to an error or owing to a holiday, we use the previous data point. We also make sure that all the data points are consistent for all instruments. For example, if there exists a holiday in Germany and the market is closed, we shall carry over the previous day's value for the German market so that it stays in tandem with the rest of the markets at the next trading day. Data that cannot be corrected is excluded from our model.

- **Jumps in data:** Where there are misprints we run a small script to identify large percentage changes. Where changes are larger than four times the standard deviation of the data series, we investigate whether this data jump is genuine or a misprint. If there is a misprint we find the correct data point from the exchange website and correct them.

- **Consistency in data and survivorship bias:** Data selection is always difficult and one needs to be careful what data to use as survivorship bias can creep in. Survivorship bias is most prevalent when one uses **all** the constituents of an index. If this index is rebalanced, with some companies leaving and others replacing them, then the index members change, bringing in survivorship bias. For example the FTSE 100 members today are not the same as one or two years ago.

  Large well-established, multi-billion pound companies, that are major employers and substantial contributors to the national GDP, such as Prudential, HSBC, Barclays have high representation (weighting) in an index such as FTSE 100. Typically, such companies do not get impacted in an index reshuffle. The key reason for this is that these companies already have high representation (weighting) in an index such as FTSE 100. Indices such as FTSE 100 remove the smaller under-performing companies, not well established multi-billion pound multinationals. To minimise survivorship bias, for our research we only focus on companies that are well-established and have large market capitalisation i.e. several billion pounds and are present in both in-sample and out-of-sample data sets. In a data set that is constantly changing, survivorship bias can be minimised, as we did, but not totally eliminated.

8.2 Model Parameters for the Four Quantitative Investment Models

We now the present the results of the four quantitative investment models from Section 5.1 using the data that we presented at the start of the chapter. As we mentioned there is an in-sample data set (2005 – 2009) as well as an out-of-sample data set (2010 – 2012). The in-sample data set is used to identify parameters and refine the model, where the objective is to maximise the returns and Sharpe ratio.

8.2.1 Parameters of the Momentum Model for Equity Indices

To build the model we use regression analysis shown in Section 3.1.3. We found that there is predictability in returns up to the 5 lag or 5 months with corresponding t-statistic. In view of conserving data in the in-sample period, to begin with we use the 5 month (approximately 107-110 working days) data points as our reference point for our moving average window size. We now search for the window size around this 5 month window size where all 17 equity index markets are profitable and the portfolio as whole has the highest Sharpe ratio. We found that the 100 days window gave good returns for all equity indices as well as a high Sharpe ratio for the momentum model portfolio.

8.2.2 Parameters of the Market Neutral Model for Global Equity Indices

This model is applied to the global equity index futures market just as in the case of the momentum model and described in Section 5.1. The global equity index futures are divided into geographic regional blocks, primarily since co-movement is time dependent; markets that are not open in a different time zone cannot react to news or events, as well as from a practical perspective so as to ensure that actual trades can be executed, with minimal market impact and liquidity risk. For example, a pairing between markets of Hong Kong and Germany is not feasible, since there is little overlap when they are open, due to the seven hour time gap. Furthermore, an event or news release that may have occurred early in the morning for Hong Kong may not be regarded as important by the time Germany opens, missing out on the move in the German market. We now present the three regional groups that we created.

- For Asia we have Japan, Hong Kong and Australian markets,
- For Europe we have UK, Germany, France Netherlands, Italy, Spain, Norway and the broader European market,
- Finally for the Americas we have USA, Canada and Brazilian markets.

The model is explained in Section 5.1. Within this framework we are optimising for two key factors: a) maximise success rate as measured by the number of profitable trades and b) maximise average gain as measured by size of average profit. To identify the appropriate window size, we test our model on several window sizes ranging from 10 days to 60 days in increments of 5 days. Empirically, we found the 10 day window is combined with lambda (measure of speed of mean reversion) threshold set at 0.1 and standard deviation (measure of dislocation in data) threshold set at 1.5, combined give the best result on average. Furthermore, just as in the case of Gatev *et al.* we pick the best performers. Gatev *et al.* pick the top 20 pairs, we pick the top 20% of the performers. Given that we have a total of 34 pairs we choose six of the best performers from them. Just as the case of Gatev *et. al* all trades have *equal capital allocation* in the portfolio, e.g. if we have ten trades then each trade has 1/10 allocation.

8.2.3 Parameters of the Market Neutral Model for Stocks

The market neutral model for EU stocks uses the same methodology as the one used for equity index pairs discussed earlier in Section 8.2.2, albeit with a little more structure. Specifically our Market Neutral model is dollar neutral, i.e. with an equal amount of capital on the *short* and *long* side of the pair, and sector neutral (Section 2.3.2). We consider a set of 250 of the largest stocks by market capitalisation in the European Union, that are denominated in Euros, and group them into the relevant sectors and sub-sectors (Appendix O). This approach works as most companies operate under very similar macro-economic and regulatory environments and hence their markets are closely linked in some cases with cross holding (see Figure 8.1). For example, in the telecommunication sector we put all the telecommunication companies from France, Germany, Spain, Netherlands, Italy etc. in the same sector and create all possible pairs such as log(France Telecom) - log(Deutsche Telekom) (see Appendix Q and Appendix P). Failing such a categorisation stocks are removed from the test data set. We go through the same methodology as we did for index pair selection and model the spread using the Ornstein-Uhlenbeck model and use the same framework to build this model (Equation 3.3).

In our in-sample tests we find that the optimal window size is 20 days, combined with lambda (speed of mean reversion) threshold, set at 0.1 and standard deviation (measure of dislocation in data) threshold set at 1.5 give the best result on average. We identify this through empirical testing on several window sizes from 10 to 50 days. We further analysed the performance of the pairs by sector and found that the insurance and utilities sectors perform best. So just as in the case of Gatev *et al.* we pick the best performing sectors, since our model is sector neutral.

We found that pairs of stocks from that same country (e.g. EDF and GDF Suez in France) have the shortest holding period and smaller returns, as a consequence mean-reversion is faster for intra-country pairs as compared to inter-country pairs. We give equal capital allocation to each trade in the portfolio, just as in the index pairs model.

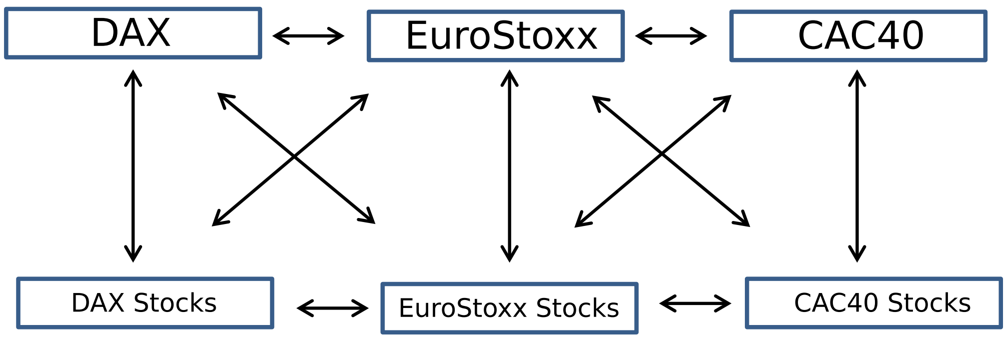


FIGURE 8.1: Stocks and indices impact each other through their interactions and working. Movements can be correlated, among indices and strongly correlated between sectors. Quantitative models can identify these relationships and turn them into investment opportunities as well as risk management tools.

8.2.4 Parameters of the Long Only Model

In the long only model we "buy" stocks and maintain a constant exposure to the market at all times. For this strategy we use both open and close prices, giving us the ability to capture more structure in data at different time horizons. The data is from FTSE 100, taking the top 100 companies by market capitalisation (equity value) and liquidity in the UK. This model is structured in a way that it buys stocks that have lagged their peers from the previous time step (see Section 5.1.3). To build this strategy we divide all the stocks into their respective sectors or sub-sectors such as banks, insurance, oil, beverages etc. We have a total of ten such sectors namely utilities, retail, banks, insurance, mining, oil and drilling, beverage, pharmaceuticals, REIT and consumer discretionary. Each sector is now a group in which the model will pick the best stock to buy (Appendix B).

8.3 In-sample Results of the Four Quantitative Investment Models

In this section we present the in-sample and out-of-sample performance of the four models, to check consistency in the performance of our models. The out-of-sample results are more important than the in-sample results. In Table 8.2 we see the performance of our individual models in the in-sample period and in Table 8.3 the correlation among the models.

The most significant thing to notice between these models is that correlation between them is low except for the two Market Neutral models that are operating at very similar time horizons and capturing similar interactions in the log spread (see Table 8.3). Both these models also have the lowest volatility of the four models. Most significantly the momentum models and the MN index model have a similar investment universe and

TABLE 8.2: In-sample performance of four underlying quantitative investment models

In-Sample	MN Index Pairs	MN Stock Pairs	Long Only	Momentum
Ann. Return	23.36%	14.68%	33.94%	44.62%
Ann. Volatility	7.10%	6.25%	17.01%	11.06%
Sharpe ratio	2.87	1.87	1.82	3.76
Prob. Gain	0.81	0.78	0.74	0.84
Prob. Loss	0.19	0.22	0.26	0.12
Avg. Gain	2.33%	1.69%	4.53%	3.84%
Avg. Loss	-0.66%	-0.74%	-3.44%	-0.98%
Gain/Loss ratio	3.56	2.28	1.31	3.91
Corr. with BAH	- 0.43	-0.26	0.86	-0.35

Key performance statistics of the four QIMs in the in-sample period. Corr. with BAH shows the correlation of the four models with the average return of all the investable assets in their investment remit. Long only model is expected to strongly correlated to its investable universe since it only buys stock. The other three models show negative correlation, showing that the models are actually exhibiting true benefit for investors as it generates returns independent of direction of the investable market.

TABLE 8.3: Correlation matrix of in-sample performance

In-Sample	MN Index pairs	MN Stock Pairs	Long Only	Momentum
MN Index pairs	1.00	0.60	- 0.24	0.19
MN Stock Pairs	0.60	1.00	- 0.30	0.20
Long Only	- 0.24	- 0.30	1.00	- 0.19
Momentum	0.19	0.20	- 0.19	1.00

Except for the market neutral models that have positive correlation on the higher side, all other models correlations are at the lower end of the spectrum. Low positive correlation is a good sign among models, and orthogonal and negative correlations are a strong plus sign for diversification of a portfolio. Most significantly the momentum models and the MN Index model have similar investment universe and exhibit very low correlation while capturing positive returns.

exhibit very low correlation while capturing positive returns. Furthermore, momentum and long only models have the highest returns, primarily because both these models are unhedged, and hence capture more of the variance than the market neutral models.

In terms of correlation with their investable universe, the market neutral models and the momentum model have negative correlation. As expected the long only model shows strong positive correlation since it only buys stocks and hence is fully exposed to the moves of the market. The momentum model shows some degree of negative correlation to its investment universe. At a time when the global markets see-sawed trending both up (positive returns) and down (negative returns) the momentum model managed to capture the momentum and be profitable with negative correlation to the market. The returns of the investment universe is the average return of all the securities, since we give equal weighting to all our trades in a model, this makes the correlation analysis relevant and interesting. The market neutral models have the lowest volatility since their investment approach has a degree of hedging, while the momentum and long only models have the highest volatility. In Figure 8.2 we can see that neither of the QIMs returns have Gaussian distribution even at monthly frequency. In the caption of the figures can see the first four moments and Jarque-Bera test statistics.

8.4 Out-of-sample Results of the Four Quantitative Investment Models

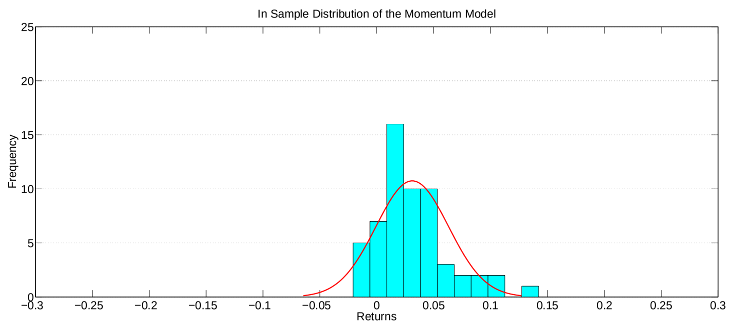
We now look at the out-of-sample results in Table 8.4 which are very similar to the in-sample results and the correlations shown in Table 8.5.

TABLE 8.4: Out-of-sample performance of the four quantitative investment models

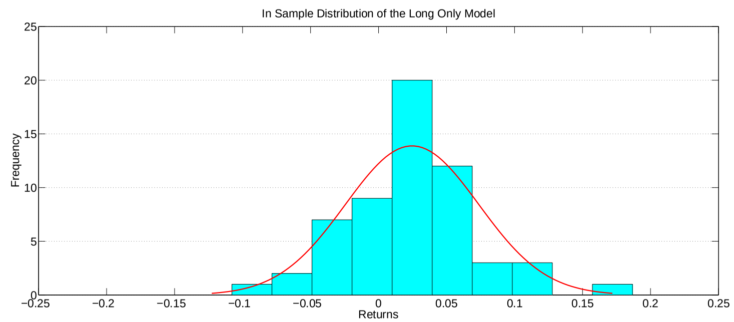
Out-of-Sample	MN Index pairs	MN Stock Pairs	Long Only	Momentum
Ann. Return	17.6%	9.9%	24.5%	46.9%
Ann. Volatility	7.72%	3.52%	14.73%	9.36%
Sharpe ratio	2.21	2.67	1.63	4.95
Prob. Gain	0.72	0.78	0.69	0.92
Prob. Loss	0.28	0.22	0.31	0.08
Avg. gain	2.33%	1.21%	4.13%	3.84%
Avg. loss	-1.18%	-0.67%	-3.34%	-3.16%
Gain/Loss ratio	1.99	1.80	1.23	1.21
Corr. with BAH	0.03	- 0.05	0.92	- 0.54

Key performance statistics of the four QIMs in the out-of-sample period. **Corr. with BAH** shows the correlation of the four models with the average return of all the investable assets in the their investment remit. Long only model just as in the in-sample period is strongly correlated to its investable universe. The Market Neutral models are almost perfectly orthogonal while the Momentum model shows negative correlation.

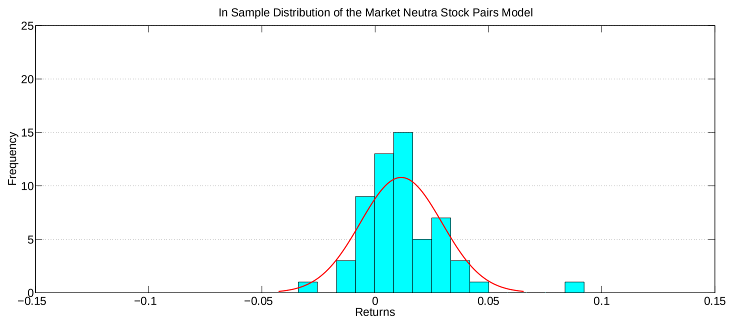
In Table 8.5 we can see that the correlation between the models is towards the lower end of the spectrum except for the market neutral models, which have slightly higher correlation but still less than the in-sample period. The momentum and the long only models have negative correlation, slightly on the higher side, which is a positive. This is good sign for our model based approach of investing on different time horizons, validating our two tier approach. The momentum model is still the best performer with a better Sharpe ratio than the in-sample period. However both the market neutral models show considerable drop in performance, except that the market neutral model for stock pairs now has a better Sharpe ratio. The momentum model still has the highest Sharpe ratio just as in the in-sample period (Table 8.4).



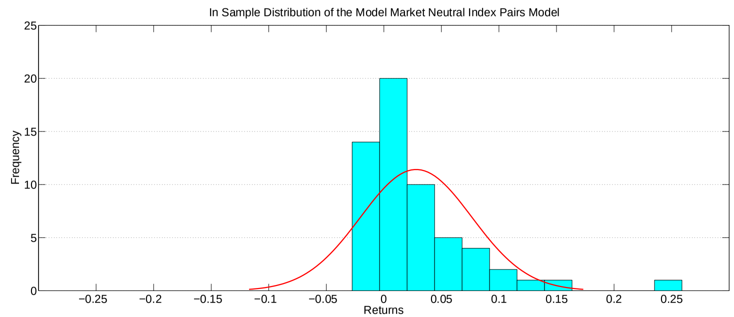
(a) In sample distribution of the Momentum Model, number of bins = 11, bin size = 0.01482, mean = 0.03122, variance = 0.00101, skew = 1.11545, kurtosis = 1.89122, Jarque-Bera statistics is 18.45137.



(b) In sample distribution of the Long Only Model, number of bins = 10, bin size = 0.02945, mean = 0.02464, variance = 0.00241, skew = 0.36463, kurtosis = 1.74318, Jarque-Bera statistics is 6.68507.



(c) In sample distribution of the Market Neutral Stock Pairs Model, number of bins = 15, bin size = 0.00840, mean = 0.01148, variance = 0.00032, skew = 1.51546, kurtosis = 6.43802, Jarque-Bera statistics is 103.34507.



(d) In sample distribution of the Market Neutral Index Pairs Model, number of bins = 12, bin size = 0.02387, mean = 0.02808, variance = 0.00234, skew = 2.36665, kurtosis = 10.71947, Jarque-Bera statistics is 198.15330.

FIGURE 8.2: In-sample distribution of QIMs. The plots are made using Freedman-Diaconis method.

Volatility of all the models has remained similar to the in-sample period, except for the market neutral model for stock pairs, which show volatility drops by half. In terms of correlation with their investment universe, the market neutral models have little to no correlation. The long only model still has strong positive correlation and the momentum model still has negative correlation to their respective investment universes. Our ability to extract positive returns from markets with low to negative correlation is a big positive for us as it shows that our QIMs add value.

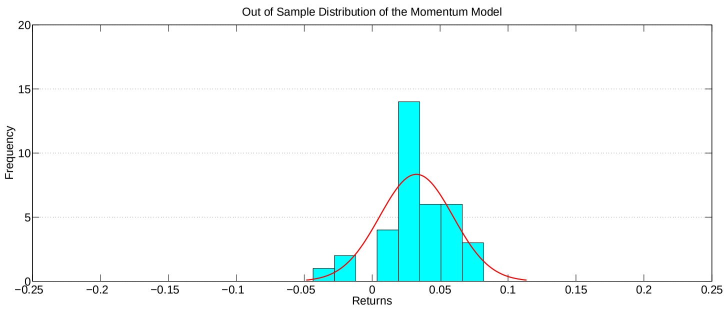
In Figure 8.3 we present the distributions of all the QIMs at monthly frequency with their corresponding first four moments and the Jarque-Bera test statistics. In Figure 8.5 we can see the capital growth of the four models in the out-of-sample period. In Figure 8.4 we can see the rolling correlation among the four models and how it evolves on monthly data. From the four charts we can see that as the window size increases the correlations are fairly stable. This stability in correlation shows that the QIMs are consistent in capturing patterns as they are designed to, even though returns in the out-of-sample period are lower than the in-sample period. In Appendix H we can see the performance of the QIMs when compared to the buy and hold strategy for their investment universe equally weighted. We can see that our QIMs outperform them.

From the QIMs that we have built, we can see from that it is possible to construct profitable models. More importantly it is possible to construct profitable models within the same asset class that capture different aspects of the data, at different window sizes and generate returns that have low correlation. Some of the correlations are low enough to make the models practically independent investments. This negates the efficient markets hypothesis which says that financial markets are efficient and process all information perfectly. However from the perspective of the Market Neutral models we can say that the markets are weakly efficient. One could also make an argument that from the perspective of market neutral models, they follow the law of one price, which roughly states that goods that are similar should have the similar value.

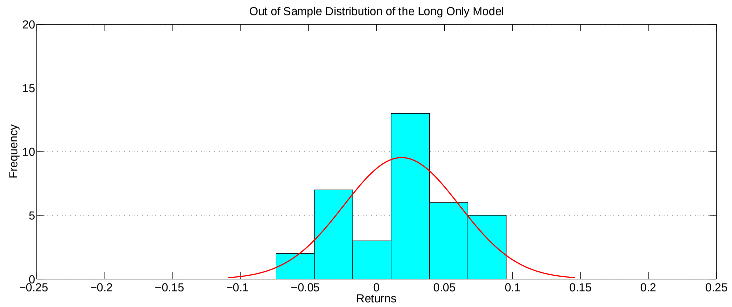
Table 8.5: Correlation matrix of Out-of-sample Performance

Out-of-Sample	MN Index Pairs	MN Stock Pairs	Long Only	Momentum
MN Index pairs	1.00	0.35	0.09	0.14
MN Stock Pairs	0.35	1.00	- 0.08	0.24
Long Only	0.09	- 0.08	1.00	- 0.55
Momentum	0.14	0.24	- 0.55	1.00

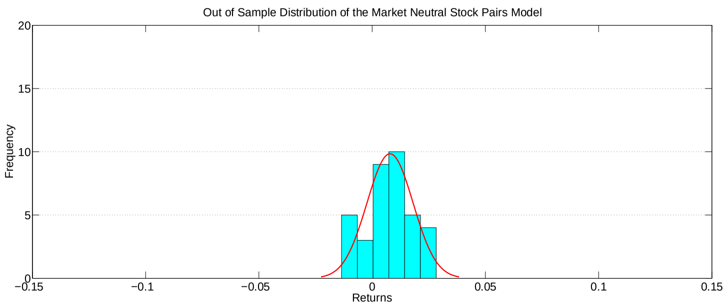
The market neutral models still have positive correlation albeit lower than the in-sample period. The momentum and long only models have stronger negative correlation in this period, which a good sign as they both generate positive returns. All other models have low correlation with each other, which is a positive sign for the meta model. Most significantly the momentum models and the MN Index model have similar investment universe and they still exhibit very low correlation again just as in the in-sample period. Models having negative correlation is a positive sign as they are all making healthy returns that are independent of other models.



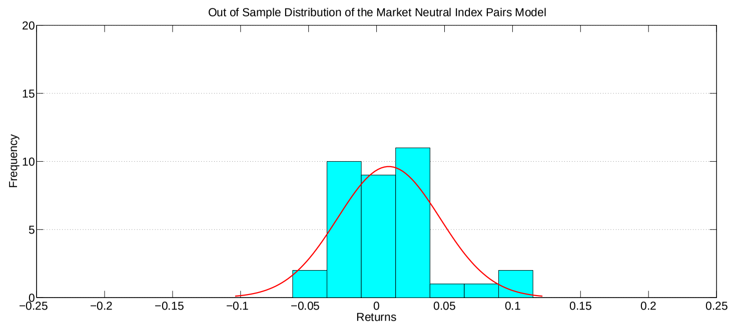
(a) Out of sample distribution of the Momentum model, number of bins = 8, bin size = 0.01569 mean = 0.03254, variance = 0.00072, skew = −0.75216, kurtosis = 1.36680, Jarque-Bera statistic = 4.96139.



(b) Out of sample distribution of the Long Only model, number of bins = 6, bin size = 0.02824 mean = 0.01844, variance = 0.00180, skew = −0.18980, kurtosis = −0.46208, Jarque-Bera statistic = 0.69080.

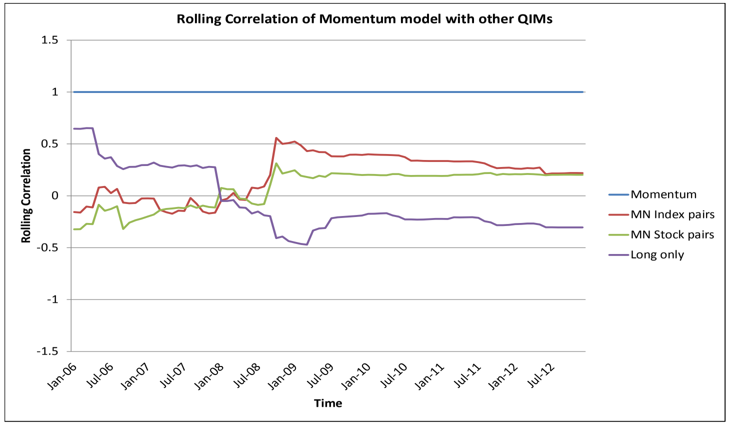


(c) Out of sample distribution of the MN Stock Pairs model, number of bins = 7, bin size = 0.00696 mean = 0.00899, variance = 0.00010, skew = −0.15275, kurtosis = −0.49803, Jarque-Bera statistic = 0.66866.

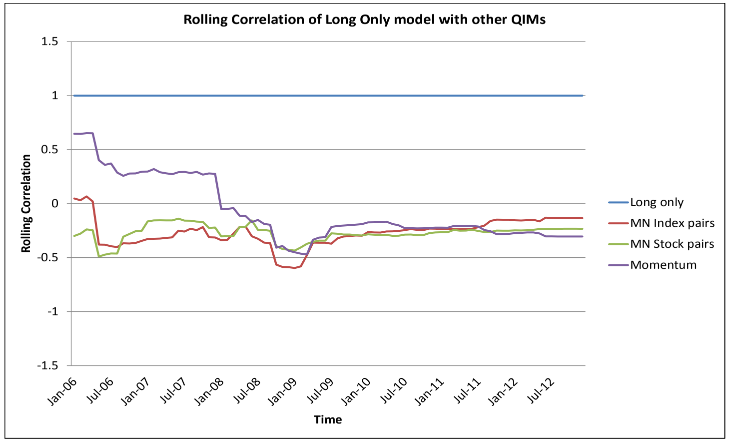


(d) Out of sample distribution of the MN Index Pairs model, number of bins = 7, bin size = 0.02524 mean = 0.00899, variance = 0.00141, skew = −0.74491, kurtosis = 1.40246, Jarque-Bera statistic = 4.99247.

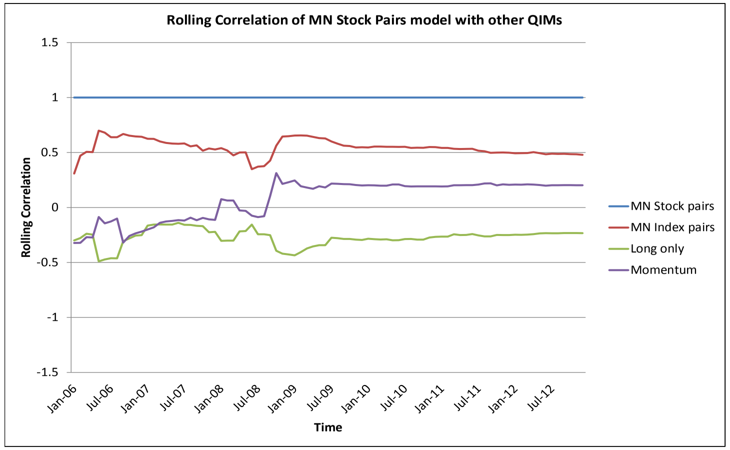
FIGURE 8.3: Out-of-Sample distribution of QIMs.The plots are made using Freedman-Diaconis method.



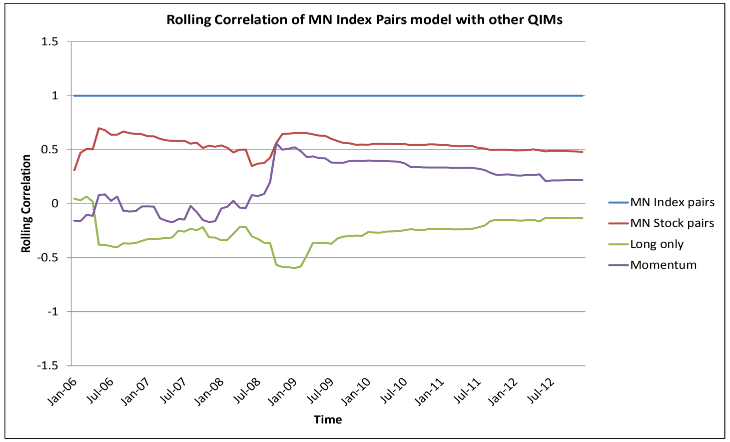
(a) Momentum model correlation with other QIMs



(b) Long Only model correlation with other QIMs



(c) MN Stock Pairs model correlation with other QIMs



(d) MN Index Pairs model correlation with other QIMs

Figure 8.4: Rolling correlation among QIMs

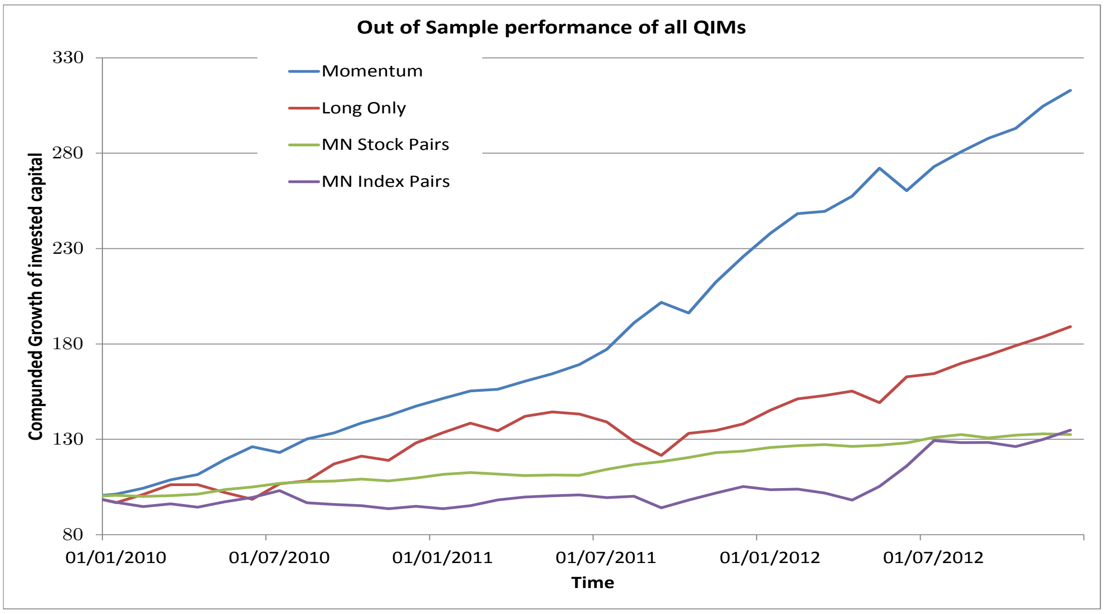


FIGURE 8.5: Performance of the four quantitative investment models in the out-of-sample period. The momentum model is the best performer.

8.5 Meta Model Portfolios of QIMs: Results

In the previous sections we presented the results of our four QIMs. In this section we present and analyse the performance of our main contribution, the meta models. To compare our meta model we use four other methods as benchmarks: a) Equally weighted (model-free), b) Mean variance optimisation framework, c) Fractional Kelly and d) Optimal Kelly. We use these methods as our benchmarks since they are appropriate for our approach, since we only have four underlying models, Portfolio by sorts is not really suitable neither is the risk parity approach, as we have already discussed in Chapter 2. We first present the in-sample performance followed by the out-of-sample performance and finally the full sample results, so we can see how our models perform. We have the distribution of all the models in Appendix K and the Kelly estimates output for all our models in Appendix L.

We evaluate the models by comparing their Sharpe ratios or risk adjusted returns, since it is a universally accepted standard for portfolio performance which incorporates both return as well as risk. For the risk-free rate we use average of the three month German government bond's yield. This became negligible in late 2010 so we have maintained it a minimum of 0.50%.

8.5.1 In-sample Results of the Meta Models

The in-sample results (2005-2009) are shown in Table 8.6 and make interesting reading. In some ways we have some observations that we would have expected, given the model while in some cases we have observations which are a good surprise. We now present some of the key performance measures below:

In Figure 8.6 we can see the annualised returns of all the meta models, Optimal Kelly is the best performing model, followed by Kelly with Kalman Filter and Kelly with Moving Average model. The rest of the Kelly models lag by a percentage point or more. Optimal Kelly being the best performer is not a surprise as it focuses on maximising return. The MVO model has the lowest returns and that is expected as this model maximises Sharpe ratio by minimising variance. Hence it eventually ends up with a portfolio that is at the lower end for returns.

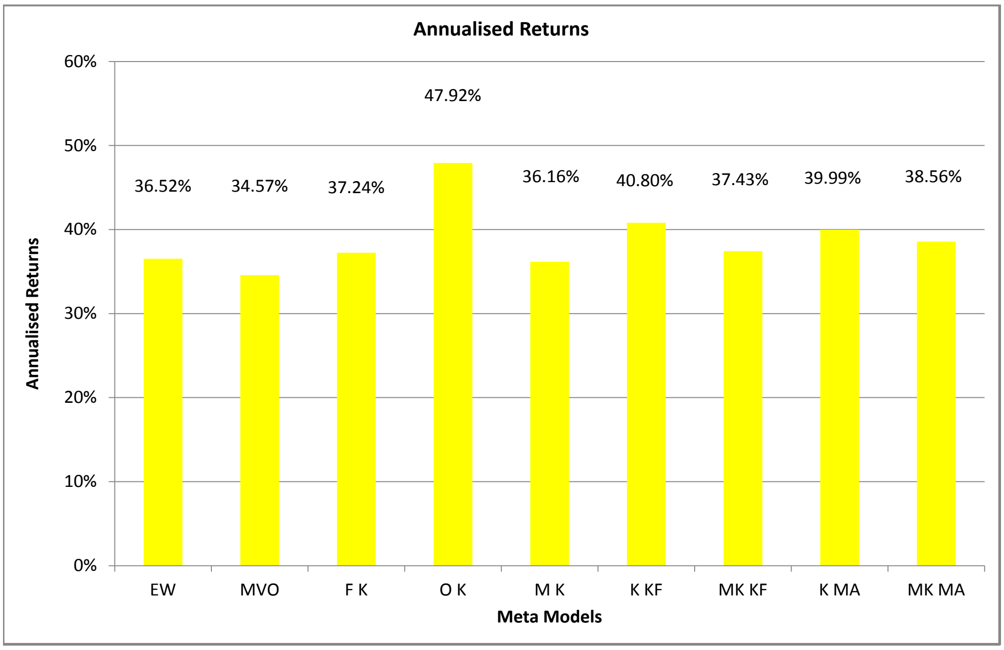


FIGURE 8.6: Annualised returns for all meta models during the in-sample period of 2005-09. Optimal Kelly is clearly the best performer of all, while MVO has the lowest returns.

In Figure 8.7 we can see the volatility of the meta models. Optimal Kelly has the highest volatility of all the models and this also impacts its Sharpe ratio. The MVO model is expected to have the lowest volatility because it captures the joint distribution. The Kelly meta models have a range of volatilities: Median Kelly with Moving Average and Kelly with Moving Average have volatilities that are on the lower end while Median Kelly with Kalman Filter and Kelly with Kalman Filter have higher volatility.

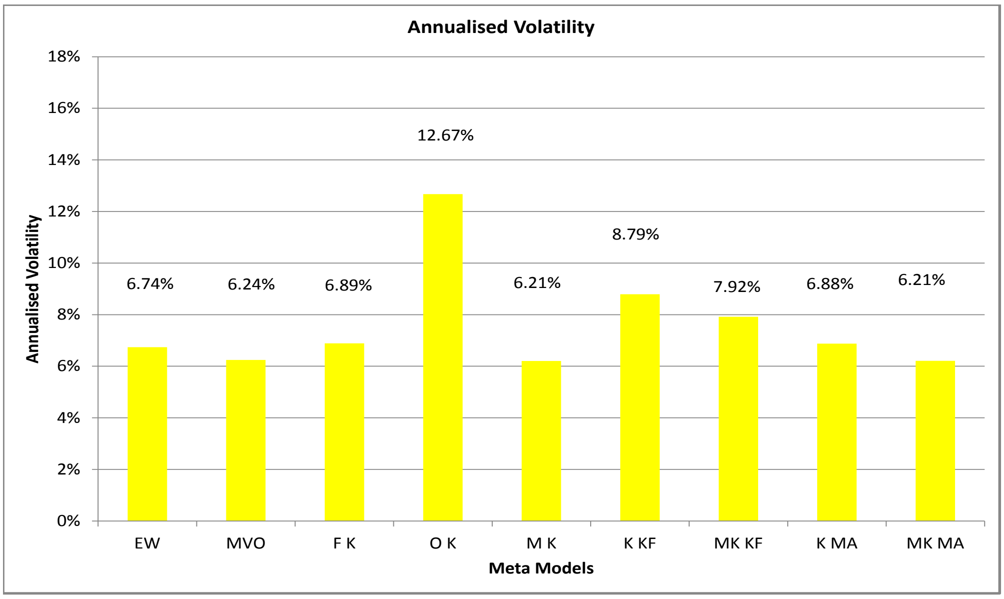


FIGURE 8.7: Volatility of all the meta models during the in-sample period 2005-09. Optimal Kelly has the highest volatility, MVO which is supposed to capture the joint distribution of returns and minimise variance is marginally beaten by Median Kelly and Median Kelly with Moving Average.

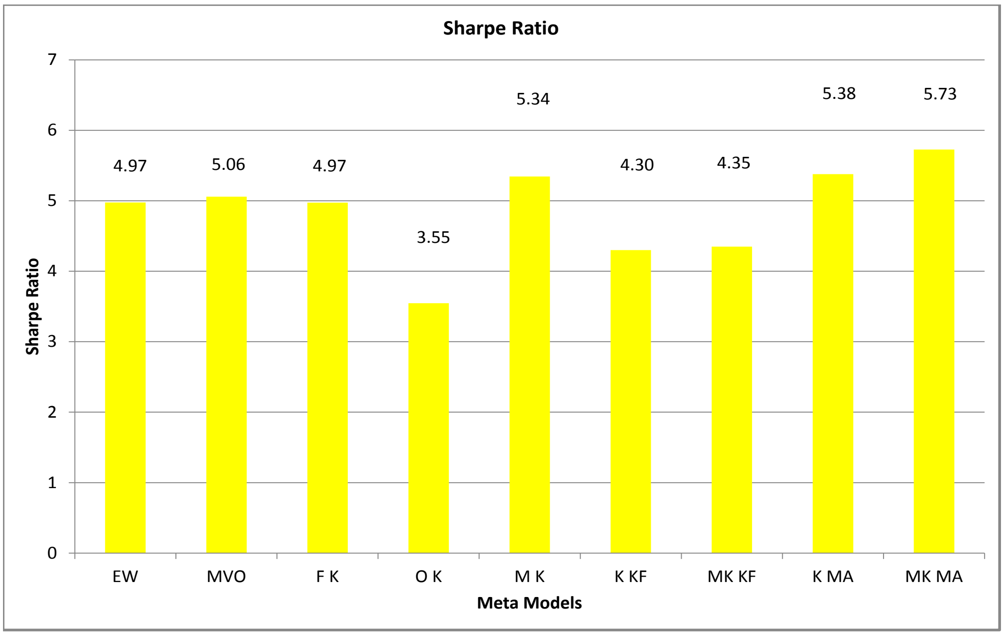


FIGURE 8.8: Sharpe ratio of all meta models during the in-sample period 2005-09. Kelly with Moving Average and Median Kelly with Moving Average are the best performers, beating all benchmarks.

Table 8.6: Comparison of meta models in the in-sample period

In-Sample	EW EW	MVO MVO	Fractional Kelly	Optimal Kelly	Median Kelly	K KF K KF	MK KF MK KF	K MA K MA	MK MA MK MA
Ann. Return	36.52%	34.57%	37.24%	47.92%	36.16%	40.80%	37.43%	39.99%	38.56%
Ann. Volatility	6.74%	6.24%	6.89%	12.67%	6.21%	8.79%	7.92%	6.88%	6.21%
Sharpe ratio	4.97	5.06	4.97	3.55	5.34	4.30	4.35	5.38	5.73
Prob of gain	0.96	0.96	0.98	0.84	0.96	0.96	0.98	0.98	0.98
Prob. Of loss	4.00%	4.00%	2.00%	16.00%	4.00%	4.00%	2.00%	0.02	0.02
Avg. win	2.74%	2.63%	2.73%	4.28%	2.72%	3.05%	2.78%	2.90%	2.82%
Avg. loss	-0.16%	-0.39%	-0.23%	-1.74%	-0.09%	-1.00%	-1.89%	-0.19%	-0.28%
Gain / Loss ratio	17.55	6.65	11.83	2.46	31.92	3.06	1.47	15.41	9.89
Median	2.36%	2.02%	2.29%	3.03%	2.17%	2.46%	2.47%	2.47%	2.58%
Average gain	2.63%	2.51%	2.67%	3.32%	2.61%	2.89%	2.69%	2.84%	2.75%

Key performance statistics for the meta models, including our four benchmarks. We can see the Kelly models perform well.

In Figure 8.8 we can see our key measure of success for a model, the Sharpe ratio. Median Kelly with Moving Average and Kelly with Moving Average have the highest Sharpe ratio, beating the benchmarks just as they did with simulated data. The Median Kelly model also does well and beats all the benchmarks. Kelly models with Kalman Filters record the lowest Sharpe ratios. This was not expected. Optimal Kelly records the lowest Sharpe ratio. In Table 8.6 we can see that the Median Kelly with Moving Average, Kelly with Moving Average, Fractional Kelly and Median Kelly with Kalman Filter have the highest probability of success. However Median Kelly has the highest gain-to-loss ratio of all the models.

Th results in the in-sample period are encouraging and validate some of the results we got from our simulated data, particularly in the case of Median Kelly with Moving Average and Kelly with Moving Average which have the highest Sharpe ratio (Figure 7.8). However, we should view these results with some scepticism as these are in-sample results.

8.5.2 Out-of-sample Results of the Meta Models

In this section we present the out-of-sample (2010-2012) results of our meta models. In Table 8.7 we show the key performance statistics. In Figure 8.9 we can see the returns of all the meta models. Just as in the case of in-sample annualised returns, Optimal Kelly is still the best performer by some margin. However, both Median Kelly with Moving Average and Kelly with Moving Average are the next best performers, beating all the other Kelly models except Optimal Kelly. They are also better than Kelly with Kalman Filter, which was the second best performer in the in-sample period.

The volatility of the meta models also makes for interesting reading. MVO, as expected has the lowest volatility; however, except for Optimal Kelly rest of the Kelly models are not very far from MVO. This is very encouraging especially in the case of Median Kelly with Moving average and Kelly with Moving Average as both the models are out-performing MVO in returns by a considerable margin. The Sharpe ratio is our key metric for performance. Both Median Kelly with Moving Average and Kelly with Moving Average beat all the benchmarks by a considerable margin. This is consistent with the in-sample results as well as the synthetic data, but as it is by a higher margin, which is a big positive for the model. In Table 8.7 we can see that the Median Kelly with Moving Average and Kelly with Moving Average also have the highest probability of success. Kelly with Moving average has the highest gain-to-loss ratio of all the models.

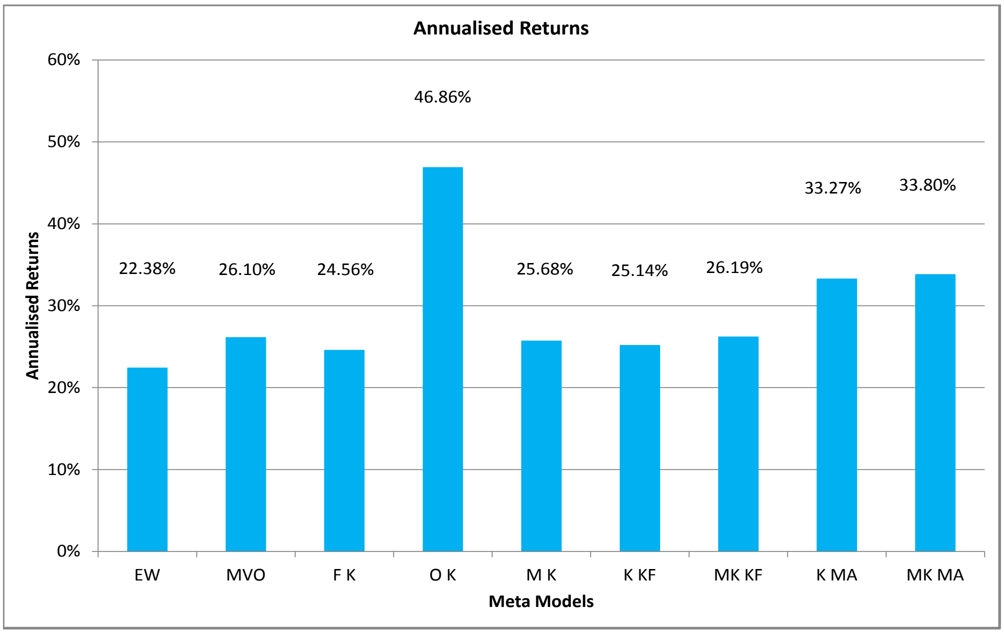


FIGURE 8.9: Out-of-sample annualised returns for all meta models. As in the simulations and in-sample period, Optimal Kelly has the best returns, followed closely by Kelly with Moving Average and Median Kelly with Moving Average, both of which beat the rest of the models by a considerable margin.

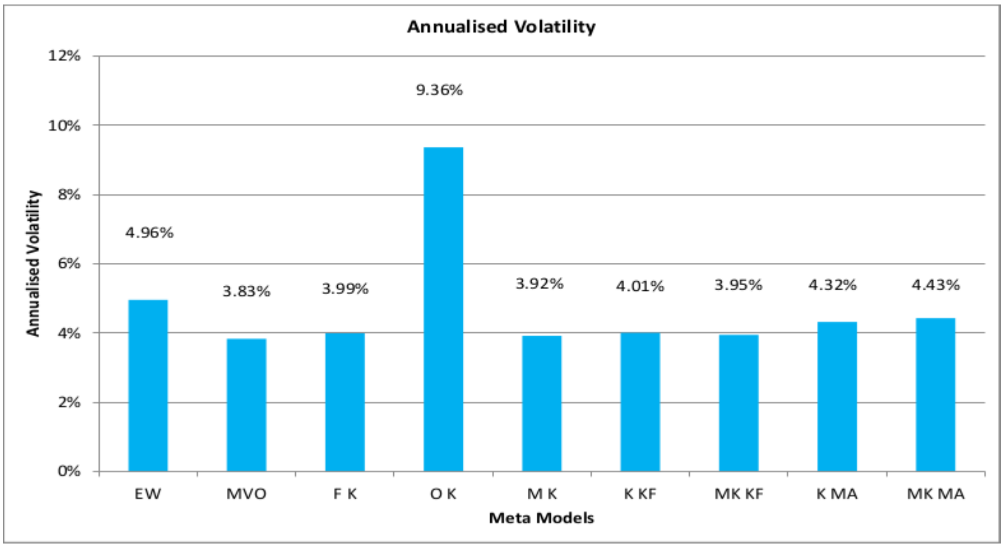


FIGURE 8.10: Out-of-sample annualised volatility. As expected, MVO has the lowest volatility, very closely followed by Median Kelly and Median Kelly with Kalman filter.

TABLE 8.7: Out-of-sample meta model comparison

Out-of-Sample	EW EW	MVO MVO	Fractional Kelly	Optimal Kelly	Median Kelly	K KF K KF	MK KF MK KF	K MA K MA	MK MA MK MA
Ann. Return	22.38%	26.10%	24.56%	46.86%	25.68%	25.14%	26.19%	33.27%	33.80%
Ann. Volatility	4.96%	3.83%	3.99%	9.36%	3.92%	4.01%	3.95%	4.32%	4.43%
Sharpe ratio	4.41	6.68	6.03	4.95	6.43	6.15	6.51	7.59	7.52
Prob of gain	0.92	0.94	0.94	0.92	0.94	0.94	0.94	0.97	0.97
Prob. Of Loss	0.08	0.06	0.06	0.08	0.06	0.06	5.56%	0.03	0.03
Avg. win	1.94%	2.09%	1.96%	3.84%	2.04%	2.00%	2.08%	2.49%	2.53%
Avg. loss	-0.97%	-0.47%	-0.13%	-3.16%	-0.14%	-0.10%	-0.11%	-0.03%	-0.16%
Gain/Loss ratio	2.00	4.50	15.00	1.21	14.43	20.44	18.11	95.84	15.54
Median	1.88%	1.95%	1.78%	3.03%	1.98%	1.81%	1.97%	2.46%	2.55%
Average gain	1.70%	1.95%	1.85%	3.25%	1.92%	1.89%	1.96%	2.42%	2.46%

Key performance numbers in the out-of-sample data set for the meta model, including our four benchmarks.

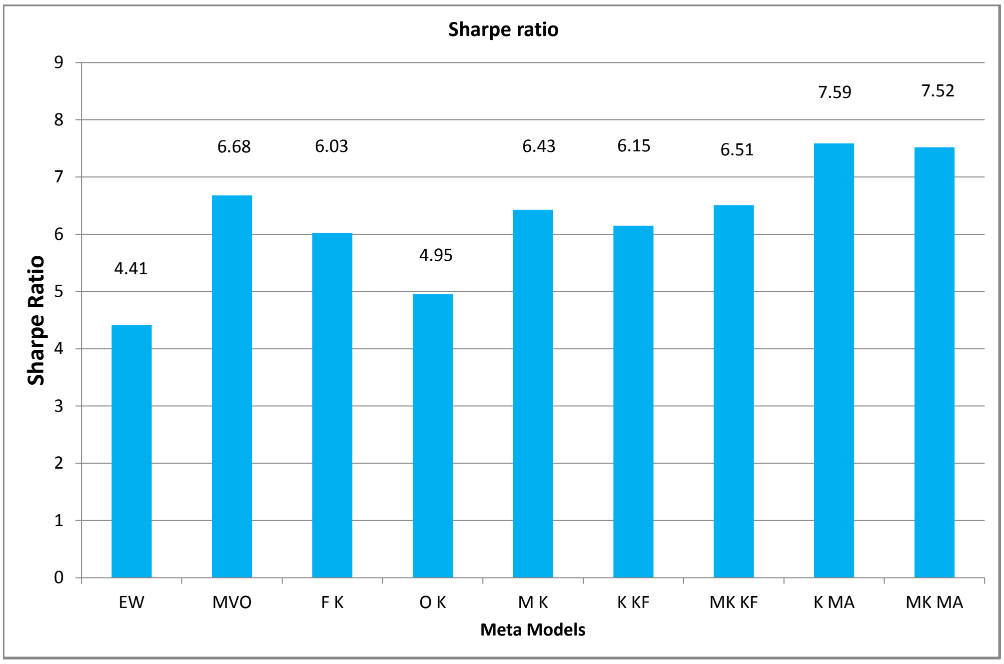


FIGURE 8.11: Out-of-sample Sharpe ratio. Kelly with Moving Average and Median Kelly with Moving Average have the the highest Sharpe ratio, in line with simulations as well as the in-sample tests, showing that Kelly with Moving Average and Median Kelly with Moving Average are better at generating risk adjusted returns.

8.5.3 Full Sample Results of the Meta Models

In the previous two sections we saw performance of the meta models both in the in-sample as well as the out-of-sample periods. However that performance is a snapshot in time and can give a skewed view. We will look at the key performance numbers as they evolve through time at monthly frequency. We will calculate the key performance measures from the in-sample period and we increase the window size till the end of the out-of-sample period, keeping the starting point static. This will show if the models perform better on a continuous and consistent basis.

We can see in Figure 8.12 that the meta model with the highest returns is Optimal Kelly, followed by Median Kelly with Moving Average and Kelly with Moving Average. The dashed black lines represent the benchmarks. This is very good for Median Kelly with Moving Average and Kelly with Moving Average models as they also have the best Sharpe ratio, as can be seen in Figure 8.13. Both these models beat the benchmarks and the rest of the models throughout.

In terms of volatility as we can see in Figure 8.14 that Optimal Kelly has the highest volatility. The lowest volatility is recorded by Median Kelly with Moving Average, Median Kelly and MVO. We expected MVO to have the lowest volatility as that is its strong point. However, we are pleased to note that both Median Kelly with Moving

Average and Median Kelly models do better than the MVO models nearly all the time except on four occasions. Since the QIMs only have week positive correlation, MVO is unable to capitalise on its strength. In Appendix F we have the key charts and performance statistics as a snapshot of the full data set for further reference.

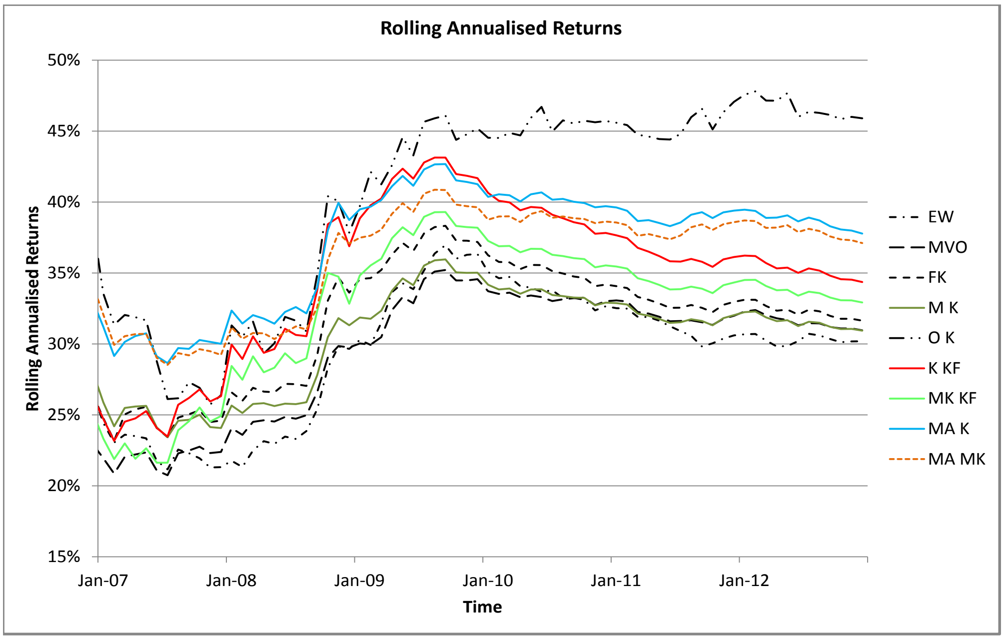


FIGURE 8.12: In this figure we can see the annualised returns of the meta models including the benchmarks (in black dashed lines) and how they evolve over time. Optimal Kelly is the best performer followed by Kelly with Moving Average and Median Kelly with Moving Average, all of them beating the benchmarks by a considerable margin.

8.5.4 Discussion: Analysis of Our Meta Models

In this section we highlight some of the key insights into our work, the pros and cons of some of the methods, as well as some important observations made during our research. MVO has been a very resilient method for constructing portfolios. Its main strength has been minimising the joint distribution of correlated random variables that have a high degree of noise. However when the data lacks strong positive correlation, the benefit of using MVO diminishes markedly, as there is little benefit or improvement to the Sharpe ratio as we see with our QIMs correlation matrix (Table 8.5). Hence, our best performing meta models (Median Kelly with Moving Average and Kelly with Moving Average) do better in terms of Sharpe ratio. Nevertheless, correlation exists in financial data whether it is stock indices, bonds, stocks or commodities, making it very hard to beat the variance minimisation properties that help MVO achieve a high Sharpe ratio.

We are able to extract some advantage by using the median rather than the mean to calculate Kelly. This results in an improvement in terms of volatility and returns, as well as the Sharpe ratio for the Kelly models. The median can be more resilient than the

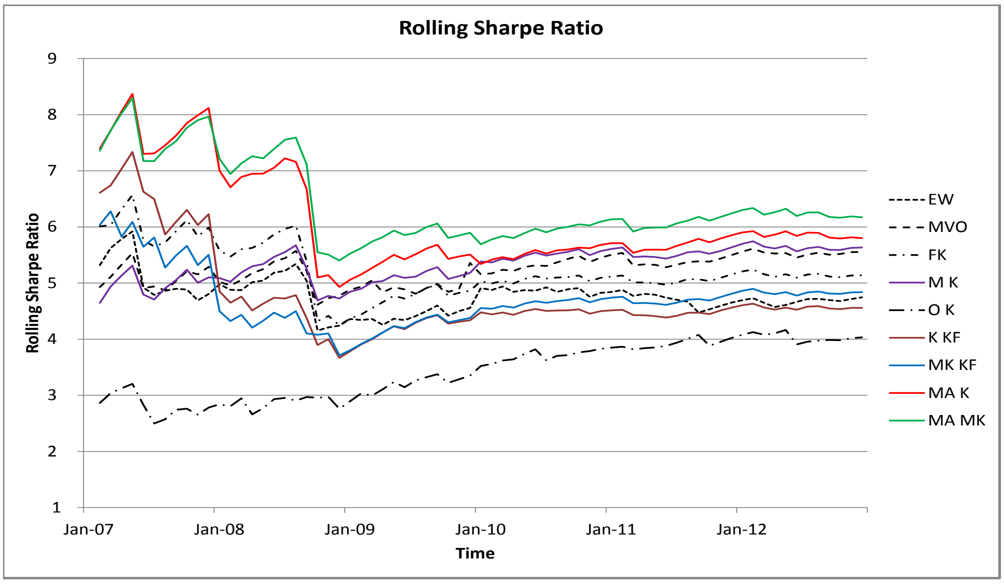


FIGURE 8.13: In this figure we show the annualised Sharpe ratio of the meta models including the benchmarks (in black dashed lines) and how they evolve over time. Median Kelly with Moving Average, Kelly with Moving Average and Median Kelly are the best peformers, outperforming the benchmarks by some margin consistently.

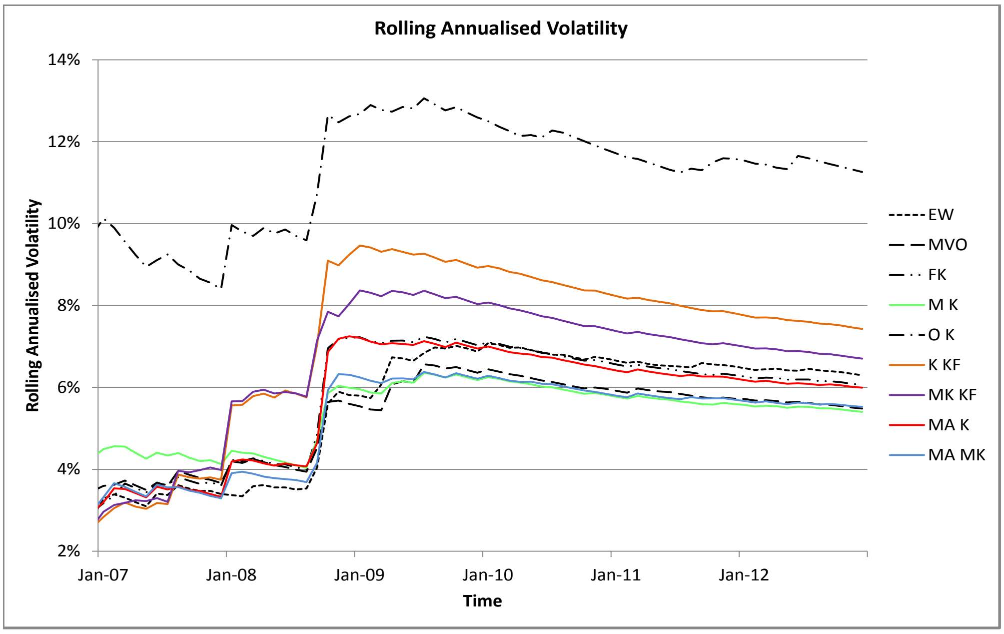


FIGURE 8.14: In this figure we show the annualised volatility of the meta models including the benchmarks (in black dashed lines) and how they evolve over time. Median Kelly with Moving Average, Median Kelly and MVO have the lowest volatility. Optimal Kelly consistently has the highest volatility.

mean and we can see this in the comparison between Sharpe ratio of Fractional Kelly vs. Median Kelly models; Kelly with Kalman Filter vs. Median Kelly with Kalman Filter; and Kelly with Moving Average vs. Median Kelly with Moving Average.

One of the key properties of Kelly is that it is myopic; i.e. to maximise log of wealth the investor only needs to know the current best investment opportunity. We wanted to use this property which was particularly applicable to our two-tier framework. A sudden change in trend or volatility could cause any of our QIMs to start making a loss. If we can avoid the loss-making QIM we would actually allocate capital in a smarter and efficient manner. We started with a Kalman Filter but during periods of introspection and discussion we realised that it wasn't the best choice owing to the assumption of Gaussian distribution. We then decided to use something simpler and more robust. Using a moving average filter really helps us manage capital allocations and improve performance.

Making use of the myopic property also helps us overcome one shortcoming of Kelly. Although Kelly has been shown to maximise wealth in the long run as it calculates the optimal bet size, in the short run it can underperform. We can see from the performance statistics that Median Kelly with Moving Average and Kelly with Moving Average both achieve higher Sharpe ratios owing to higher returns and not owing to lower volatility. Although the volatility is low for Median Kelly with Moving Average and Kelly with Moving Average it is not always lower by a large amount when compared to MVO. It is worth noting that the MVO method usually achieves a higher Sharp ratio by minimising volatility as it captures the joint distribution of the portfolio.

The meta model that maximises gains best is Optimal Kelly. The key problem with Optimal Kelly is that it achieves its goal of maximising wealth but it allocates all its capital to one model, i.e. the one that is going to maximise its return at that moment. The model with the highest return dominates all other models, except on two occasions in the in-sample period. We can see that Optimal Kelly allocates all its investment capital to the momentum model as its performance is identical in terms of returns, volatility and Sharpe ratio in the out-of-sample period (see Table 8.7 and Table 8.4). We present the weights for Optimal Kelly upto five decimal places in Appendix M.

Laureti et al. (2010) [1] conducted a comprehensive analysis of Optimal Kelly and MVO generated portfolios to show that the Optimal Kelly portfolio as a subset of MVO lies on the efficient frontier which is built using the MVO approach. We tested this claim with the data from our QIMs and we can confirm this is the case. In Figure 8.15 we plot the returns generated by the Optimal Kelly and the most profitable portfolio generated by the MVO model. We can see that indeed the Optimal Kelly returns are nearly identical. Just as in the case of Optimal Kelly, the highest return portfolio from MVO

[1] Conversation with Professor Doyne Farmer at IEEE Conference for Computational Intelligence for Financial Engineering and Economics 2014 regarding Kelly lead to discussion about research done by (Laureti et al., 2010) he agreed with claim made by Laureti et al.

allocates all its capital to one model that maximises returns. Furthermore we compare the annualised volatility of these returns in Figure 8.16, since they must match as well and once again we see that the volatility is the same. The slight difference in returns comes from difference in weight in the in-sample period that we pointed out above.

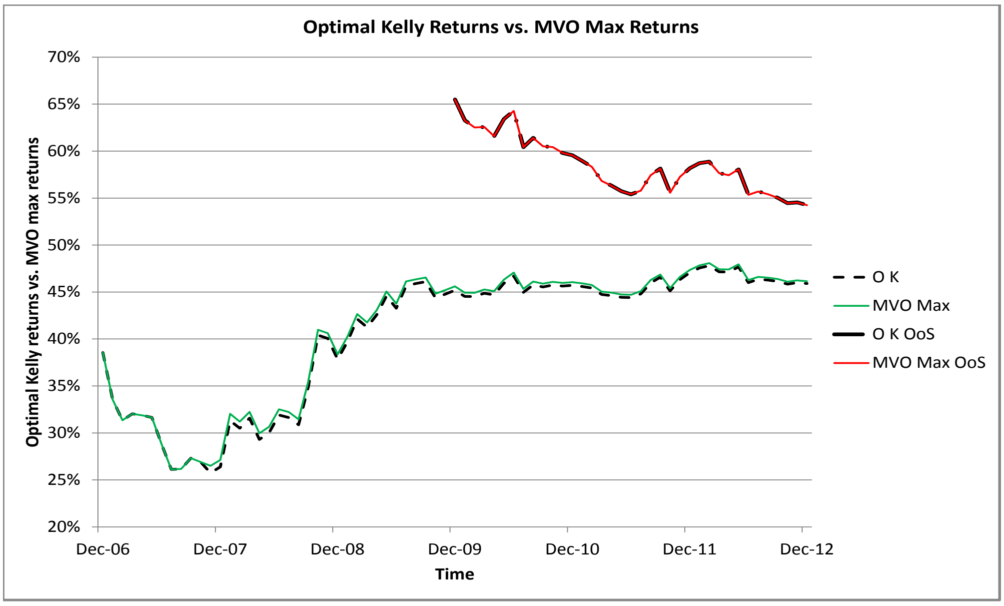


FIGURE 8.15: In this figure, we show the annualised return of Optimal Kelly compared to the annualised return of MVO when MVO picks the portfolio that maximises returns, instead of the Sharpe ratio. We show annualised returns from the in-sample period to the out-of-sample period with the green and dashed black lines. The out-of-sample period's annualised returns are shown by the red line and the dashed black line. We can see that the annualised returns of Optimal Kelly are identical to MVO max out-of-sample.

8.5.4.1 Strengths of Our Meta Model

From a computational perspective Kelly with Moving Average is much more efficient than the MVO approach. Since Kelly is myopic, the moving average only needs to know its previous position to extrapolate the future position. Both are light on memory as well as computation which is in stark contrast to MVO which needs a computationally intensive optimiser. Both Kelly and moving average make no distributional assumptions about the data, so are unburdened by assumptions; hence these models are going to be robust in the face of changing data. Kelly with Moving Average is straightforward in detail and easily generalisable. It can be applied to other domains where resource allocation is important, such as infrastructure construction, manufacturing etc. Our approach highlights that we can improve performance if the problem to be solved is specified in detail. This highlights that by combining higher order models with lower order models, one can get better results.

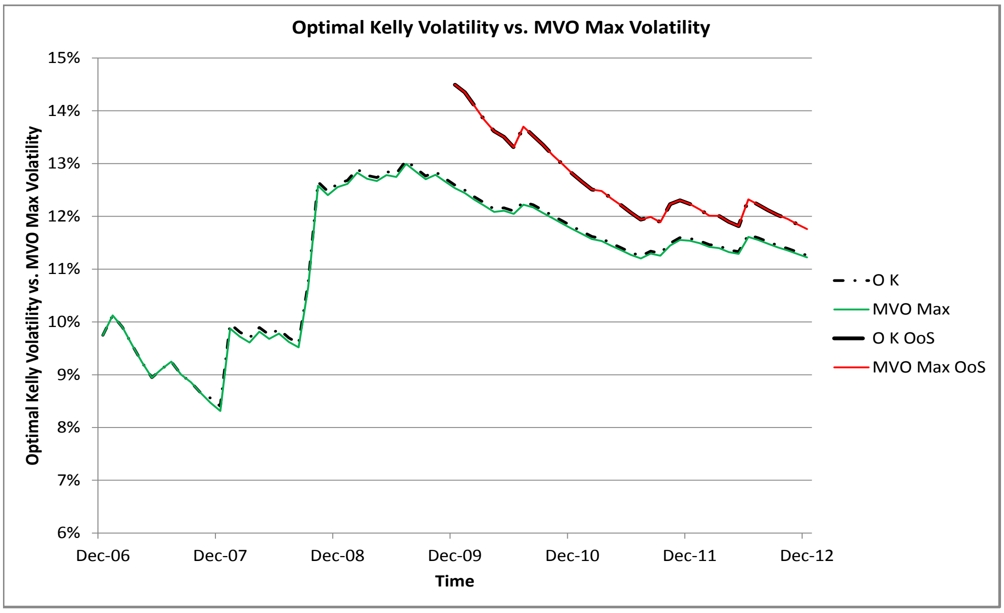


FIGURE 8.16: In this figure we show the annualised volatility of Optimal Kelly compared to the annualised volatility of MVO portfolio when MVO identifies the portfolio that maximises returns instead of the Sharpe ratio. We show the annualised volatility from the in-sample period to the out-of-sample period with the green and dashed black lines. The out-of-sample period's volatility is shown by the red line and the dashed black line. We see that the annualised volatility of Optimal Kelly is identical to MVO Max out-of-sample.

One of the key points that we have learnt over the years is that MVO builds an efficient frontier out of optimal portfolios which allows an investor to choose a portfolio that fits their risk-return profile or better yet their utility function. The Median Kelly with Moving Average and Kelly with Moving Average models have breached that efficient frontier, as can be seen in Figure 8.13 and the results from the synthetic data in Table 7.2. Both these models consistently beat that optimal portfolio as measured by the Sharpe ratio, which is part of the efficient frontier. This is a significant result since we have thought of optimal portfolios as efficient portfolios. The best portfolios all exist somewhere on the efficient frontier, except in our two-tier framework.

The Moving Average filter that we apply to avoid investing in loss-making models also helps against models that might have errors or are misspecified, i.e. if they become loss-making then the filter will stop you from investing in that model.

8.5.4.2 Shortcomings of Our Approach

We believe one of the shortcoming of our approach was that we did not have more QIMs, to capture more patterns and dimensions in the data. Part of the problem was that we didn't have access to much more expensive fundamental (accounting) data for stocks, which would have potentially given us another model with low correlation to other models. We would have also benefited from longer history for all our data sets which would have allowed us to build longer-term mean-reversion models for the equity market, the kind that generally occurs over a period of three to four years and are dependent on the economic cycle.

8.6 Summary

In this chapter we presented the performance results for our four QIMs as well as our meta models, using real FTS data. We found the models were profitable in both the in-sample and out-of-sample period, although there was a drop in performance of some QIMs. We then presented the performance of our meta models and their benchmarks. In the in-sample period and the out-of-sample period, Kelly with Moving Average and Median Kelly with Moving Average beat the benchmarks. This is in line with the tests done with synthetic data in the previous chapter. We then checked the performance of the meta models on a continuous basis, and there too we find that Kelly with Moving Average and Median Kelly with Moving Average beat the benchmarks. We also checked if Optimal Kelly was a special case of MVO as shown by Laureti et al. (2010). We found that indeed was the case.

One of the key points to note is our two-tiered approach, where the first tier focuses on certain patterns in data at a specific time horizon, with a dedicated QIM and second tier, which is the meta model that allocates capital, forms a meta portfolio. All the meta models have better Sharpe ratio than our four QIMs. This is a significant enhancement in performance and strong validation of our two-tier framework. We believe this is an important point that we are able to achieve better results through this approach. We also believe that by adding more models to the data at different time horizons we can achieve better results and use the data even better at different time horizons.

We believe both Kelly with Moving Average and Median Kelly with Moving Average beating the benchmarks can be generalised to other problems especially in cases where the data is heterogeneous as Kelly makes no distributional assumptions. As discussed in Chapter 1 we built a set of models that give us investment opportunities with high probability of success, calculate optimal investment size and control for risk. All of these are important for prudent portfolio management.

Chapter 9

Conclusion & Future Work

In this thesis we presented a smart portfolio management framework that targeted patterns found in financial data at different time horizons, while making smart and optimal bets using QIMs. To capture structure and pattern in data at different time horizons we built four QIMs focusing on the equity markets, two market neutral models using pairs methodology, and two models targeting momentum, using both stocks and index futures. The QIMs gave us high probability investment opportunities in the markets. The framework then constructed a meta portfolio by making optimal investments using Kelly Criterion. We chose Kelly Criterion as it has been shown to compute optimal investment size.

We developed several models of which we presented five key ones that show the evolution of our work. Median Kelly model gave us good Sharpe ratio measures. However regimes in financial time series can change over a period of time. This has implications for a model, as it may stop being profitable once the regime it is targeting has stopped or evolved into another one over a period of time. This is a real and tricky challenge any smart investor or portfolio manager would ideally like to avoid. To overcome this challenge we devised a Kelly model using moving average as filter to avoid investing in a loss-making QIM. We tested the models with synthetic data using a GARCH model and real financial data. The Kelly models with moving average had consistent performance, beating the efficient frontier consistently.

Throughout our research we were driven by the thought that if we broke down the challenge into its important parts and specified the challenge correctly we could find a solution to the challenge. We are pleased we managed to do that with our two-tier framework. We believe the framework, as well as the models, open up a path to using financial data in an encompassing manner and hence managing a portfolio in a smarter manner. Our meta models especially Median Kelly with Moving Average and Kelly with Moving Average are generalisable and can be applied to similar challenges in other domains for resource allocation, such as infrastructure construction and manufacturing.

9.1 What are the Implications?

We hope our research opens up this line of inquiry further and more researchers explore this path and use bigger and comprehensive data sets to test the meta model. The implication for asset managers are significant. They should heed some of the advise here and start to look at their investment remit and add models that capture patterns in data that can be turned into portfolios for investors. By not doing so, they are at the very least being inefficient and potentially managing portfolios in a sub-optimal manner.

Our model is suited for risk-averse investors who want to make smart investments yet protect themselves from losing money in the long run. The kind of investor it suits is the average man who has a modest pension portfolio and is risk averse The downside is that the average person would need some technical skill to use this model. A lack of appreciation for structure or patterns in data at different time horizons clearly shows that they are missing some opportunities. On the other hand, if once they have a lower level model ready they will be able to improve the Sharpe ratio of their portfolios by using our meta models.

Our meta model will certainly help reduce risk but it has its limitations. Not everybody can use the model at the same time as investing is a zero sum game; i.e. for each buyer there is a seller. Depending on how the price moves there is always a winner and a loser, so not every one can use this approach at the same time. Our model is also limited by the amount of capital it can take; for example the model would struggle to manage very large amounts of capital, such as above £5 billion even for a truly global portfolio. Large pension funds manage well above £100 billion. They would only be able to manage a fraction of their portfolio in this manner. Nevertheless using our approach some portion of the portfolio would be a smart portfolio.

9.2 Future Work

One of the key shortcomings for us was the lack of a comprehensive data set. By comprehensive we mean with a longer history, accounting data, observations at different frequencies, such as hourly, minute and tick data. A longer history would have helped us capture mean-reversion on much lower frequency such as five to seven years, as business and interest rate cycles impact certain sectors and markets, significantly offering investment opportunities. Furthermore accounting data would help us capture a different aspect of financial data through multi-factor models used by large institutions. Incorporating transaction costs combined with liquidity and execution risk would bring the model closer to real markets. Availability of higher frequency data would have most certainly helped us capture more patterns such as hourly and minute-by-minute momentum, as well as mean-reversion. We would also like to add further models that capture a different aspect of the data. For example, where possible we should develop models to target implied volatility using options, convertible bonds and warrants. Furthermore we would like pursue methods that look at the whole capital structure of companies i.e. using both equity and debt to capture yield differences.

A natural extension of our approach, which has only used the equity market so far, would be to expand our markets and include foreign exchange and bonds, both corporate and government. This expansion into different asset classes will broaden our use of the risk profile of the framework, as well as give us more data from different markets, making our approach comprehensive.

Appendix A

Definitions

In this section we define terms that will be used throughout the report. We aim to cover relevant definitions in statistics and finance. For statistics we cite the Cambridge Dictionary of Statistics 2006

- **Autoregressive Conditionally Heteroscedastic (ARCH):** ARCH is a class of models that are used to model data that has heteroscedasticity(if variance of the errors is not constant this would be known as heteroscedasticity). ARCH models are also used in scenarios that have volatility clustering, the tendency of large changes in asset prices (of either sign) to follow large changes and small changes (of either sign) to follow small changes. In other words, the current level of volatility tends to be positively correlated with its level during the immediately preceding periods. (Brook pg 386-7)

- **Autocorrelation:** The internal correlation of the observations in a time series, usually expressed as a function of the time lag between observations. Also used for the correlations between points that are different distances apart in a set of spatial data (spatial autocorrelation). The autocorrelation at lag $k, \gamma(k)$, is defined mathematically as

$$\gamma(k) = \frac{E(X_t - \mu)(X_{t+k} - \mu)}{E(X_t - \mu)^2}. \tag{A.1}$$

  Where $X_{t1} = 0, \pm 1, \pm 2 \ldots$ represents the values of the series μ, is the mean of the series, E denotes expected value.

- **Beta:** Beta is that part of a stock movement which is driven by the market index, of which the stock is part of. Beta is also a regression coefficient in times series regression

- **Brownian motion:** A stochastic process, X_t, with state space the real numbers, satisfying

$$X_0 = 0. \tag{A.2}$$

for any $s_1 \leq t_1 \leq s_2 \leq t_2 \dots s_n \leq t_n$ the random variables $X_{t_1} - X_{s_1} \dots X_{t_n} - X_{s_n}$ are independent.

For any $s < t$, the random variable $X_t - X_s$ has a normal distribution with mean 0 and variance $(t-s)\sigma^2$.

- **Call option:** A Call option gives the holder the right to buy the underlying asset by a certain date for a certain price (Hull 1996)

- **Correlation:** A general term for interdependence between pairs of variables. It is measured through the correlation coefficient which is an index that quantifies the linear relationship between a pair of variables. In a bivariate normal distribution, for example, the parameter, ρ. An estimator of ρ obtained from n sample values of the two variables of interest, $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$, is *Pearson's product moment correlation coefficient, r,* given by:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}. \tag{A.3}$$

- **Covariance:** The expected value of the product of the deviations of two random variables, x and y from their respective means μ_x and μ_y, i.e.

$$cov(x, y) = E((x - \mu_x)(y - \mu_y)). \tag{A.4}$$

- **First order difference:** Is simply the value that is derived from difference of two real number $g \in \mathbb{R}$ usually ordered $g_t - g_{t-1}$.

- **Futures:** A futures contract is an agreement between two parties to buy or sell an asset at a certain time in the futures for a certain price. (Hull 1998)

- **Heterogeneous:** A term used in statistics to indicate the inequality of some quantity of interest (usually a variance) in a number of different groups.

- **Intercept:** The parameter in an equation derived from a regression analysis corresponding to the expected value of the response variable when all the explanatory variables are zero.

- **Long:** To be long is the practice of buying an asset in the expectation that it will go up in value.

- **Mean:** A measure of location or central value for a continuous variable. For a sample of observations, $x_1, x_2, \dots x_n$, the measure is calculated as

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}. \tag{A.5}$$

- **Momentum:** Momentum or trend is the continued movement of price either up or down.

  A price trend model of a price series z is defined by: *(Taylor 1980)*

$$\log(x_t) - \log(z_{t-1}) = x_t = \mu_t + e_t. \tag{A.6}$$

$$E(e_t) = 0, E(e_t - e_{t+i}) = 0, (i \neq 0)\, cov(\mu_s, e_t) = 0(alls, t). \tag{A.7}$$

- **Mean reversion:** A measure of μ is said to be invariant for (X_t) if and only if

$$\int \mu(dx)\, P_t f(x) = \int \mu(dx)\, f(x). \tag{A.8}$$

  for any bounded function f. μ is said to be invariant for (X_t) if and only if $\mu P_t = \mu$. Equivalently the law of $(X_{t+u})_{u \geq 0}$ is independent of t if we start at date 0 with the measure μ.

  In econometric terms testing for mean reversion is performed through a test for stationarity by performing the Augmented Dicky-Fuller (ADF) test. The ADF consists of estimating the regression ρ_t on ρ_{t-1}, if this coefficient is below 1, it means the price is mean reverting; if it is close to 1, the process is a random walk.

$$X_{t+1} = \rho X_t + \varepsilon_t. \tag{A.9}$$

- **Stochastic process:** A series of random variables, X_t, where t assumes values in a certain range T. In most cases x_t is an observation at time t and T is a time range. If $T = \{0, 1, 2 \ldots\}$ the process is a discrete time stochastic process and if T is a subset of the nonnegative real numbers it is a continuous time stochastic process.

- **Standard deviation:** The most commonly used measure of the spread of a set of observations. Equal to the square root of the variance.

- **Stationarity:** A term applied to time series or spatial data to describe their equilibrium behaviour. For such a series represented by the random variables, $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$, the key aspect of the term is the invariance of their joint distribution to a common translation in time. So the requirement of strict stationarity is that the joint distribution of $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$ should be identical to that of $X_{t_{1+h}}, X_{t_{2+h}}, \ldots, X_{t_{n+h}}$ for all integers n and all allowable h, $-\infty < h < \infty$.

- **Systematic risk:** Risk that is inherent in the market and cannot be diversified.

- **Short:** To be short in finance is a term used when someone wants to profit from falling prices and does not own the security. They borrow the security from a broker or custodian in the market and then sell it. At a later date they buy back the security from the market and return it to the borrower. Lenders generally charge a small fee for lending the security.

- **Stylized fact:**A fact of the real world simplified and made more abstract to be usable in an economic model. Each school of economics has its favourite stylized facts, e.g. that there are steady long-term capital output ratios and Kuznets view that the average propensity to consume is relatively constant over long periods.

- **Sharpe Ratio:** Developed Nobel Laureate Williams. The Sharpe Ratio is calculated by substracting the risk free rate from returns and divided by the standard deviation of returns. The ratio measures the relationship of reward to risk in an investment strategy.

- **Unsystematic risk:** Stock specific risk that can be diversified through portfolio construction.

- **Utility:** The satisfaction derived from an activity, particularly consumption. The total amount of such satisfaction is total utility; the satisfaction from the last unit is marginal utility. Bentham in his suggested calculus of pleasure and pain was influential in introducing this notion into economics but the marginalists were the first economists to make it the central concept of economic theory. The measurement of utility has provoked long debates between cardinal utility (utility measured in units) and ordinal utility (utility revealed through preferences). Without this concept, much of neoclassical economic theory would not be possible. Earlier economic writers, especially those of the classical school, used 'utility' in the objective sense of the inherent worth of something.(Majumdar, T. (1961) The Measurement of Utility, London: Macmillan.)

- **Utility function:** This is generally expressed in the form $U = f(x_1, x_2, x_3, x_4, x_5, \ldots)$. It shows a consumer's utility as a function of the quantities of goods and services $1, 2, 3, 4, 5, \ldots$ he or she consumes.

- **Variance:** In a population, the second moment about the mean. An unbiased estimator of the population value is provided by s^2 given by

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 . \qquad \text{(A.10)}$$

  where $x_1, x_2, \ldots\ldots x_n$ are the n sample observations and $\overline{x}$ in the sample mean.

- **Volatility:** Volatility is estimated as the standard deviation of the logarithm of the ratio of stock prices divided by the square root of the length of the time period in years. (Hull 1998)

- **Warrants:** Warrants are call options that often come into existence as a result of a bond issue. (Hull 1998)

Appendix B

Long Only Sector Groups

Table B.1: sector Grouping for Long Only Model

Sector	Company	Company	Company	Company
Utilities	United Utilities	Centrica	Scottish & Southern	National Grid
Retail	Morrisons	Marks and Spencer	Sainsbury	Tesco
Banks	HSBC	Lloyds	Royal Bank of Scotland S	Barclays
Insurance	Royal Sun Alliance	Aviva	Prudential	Standard Life
Minning	Anglo American	BHP Billiton	Rio Tinto	Antofagasta
Oil and Drilling	British Petroleum	Royal Dutch Shell	Tullow Oil	
Beverage	Diageo	SAB Miller	Smith & Nephew	
Pharmaceuticals	Astra Zeneca	Glaxo Smithkline Beecham	Shire Pharma	
REIT	British Land	Hammerson	Land Securities	
Consumer Discretionary	Assc. British Food	Whitbread	Kingfisher	

10 sector groupings from the constituents of FTSE 100, used for the Long Only model.

Appendix C

Rolling Correlation amongst QIMs

Here we graphically show the rolling correlation of QIMs.

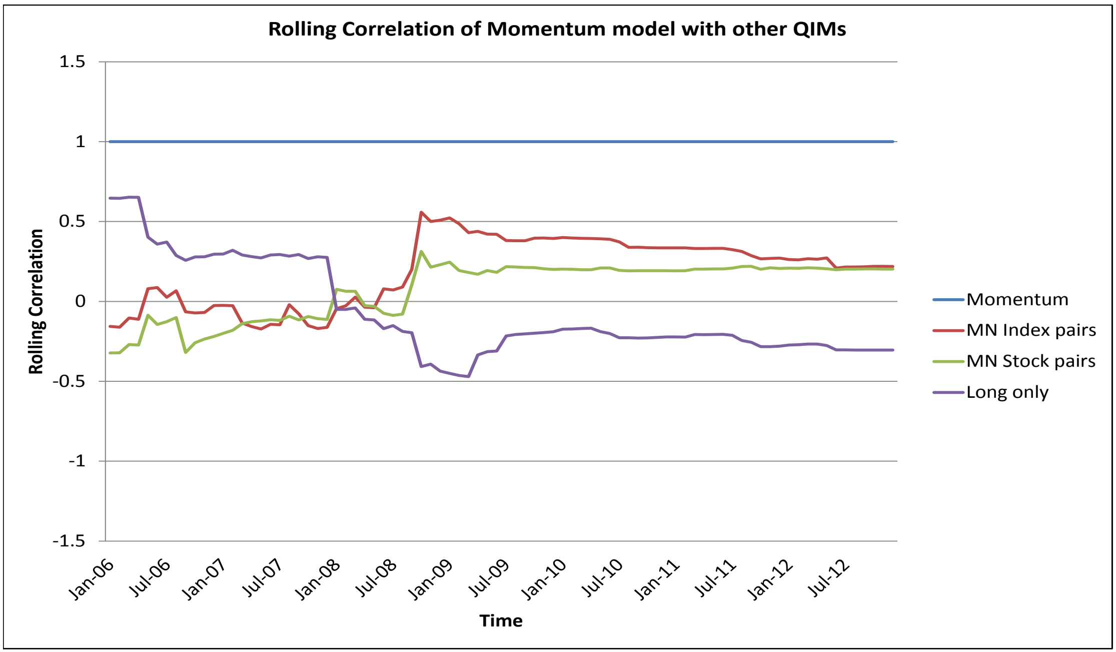


Figure C.1: Momentum QIM's continous correlation to the rest of the QIMs.

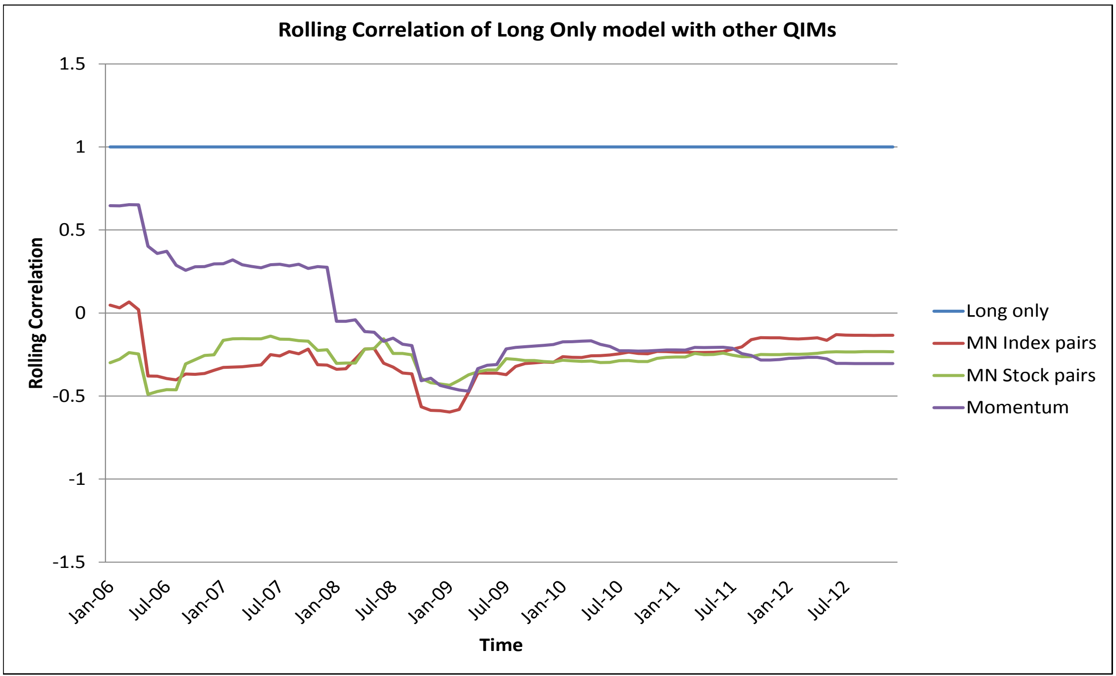


FIGURE C.2: Long Only QIM's continous correlation to the rest of the QIMs.

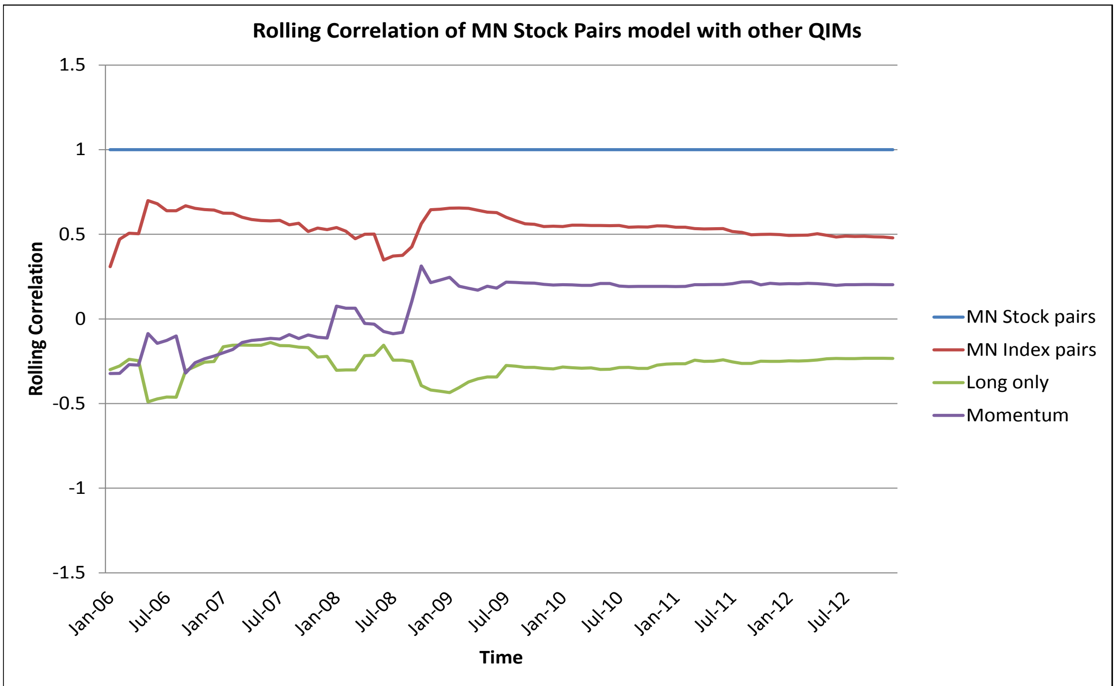


FIGURE C.3: MN Stock pairs QIM's continous correlation to the rest of the QIMs.

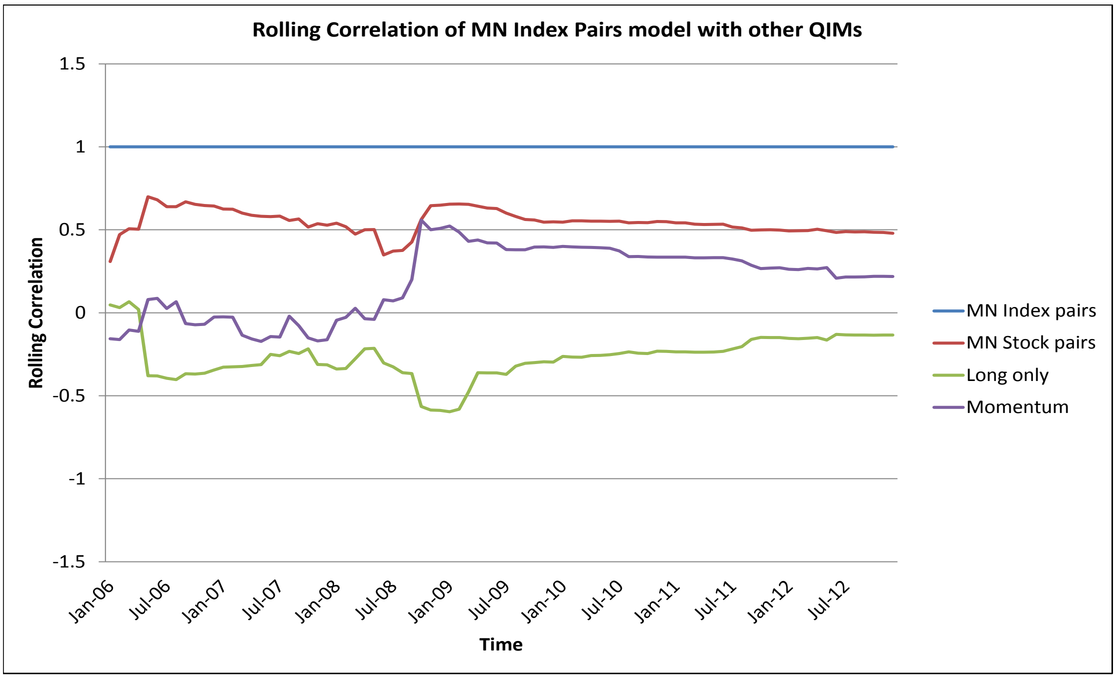


FIGURE C.4: MN Index pairs QIM's continous correlation to the rest of the QIMs.

Appendix D

In-sample meta model Performance

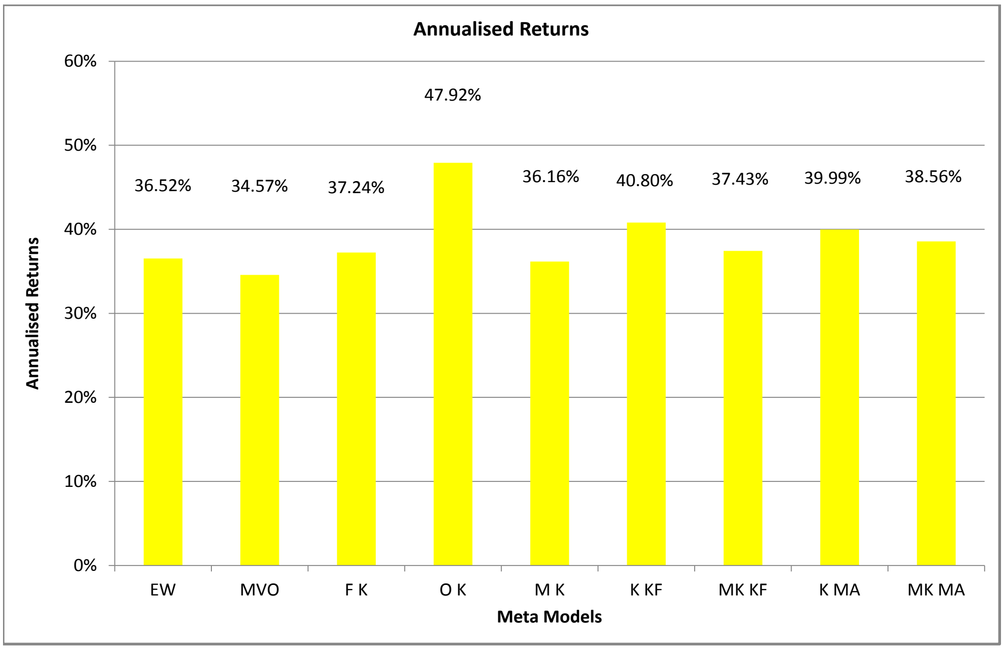


FIGURE D.1: Annualised returns for all meta models during the in-sample period of 2005-09. Optimal Kelly is clearly the bets of all, EW is the only benchmark that has high returns.

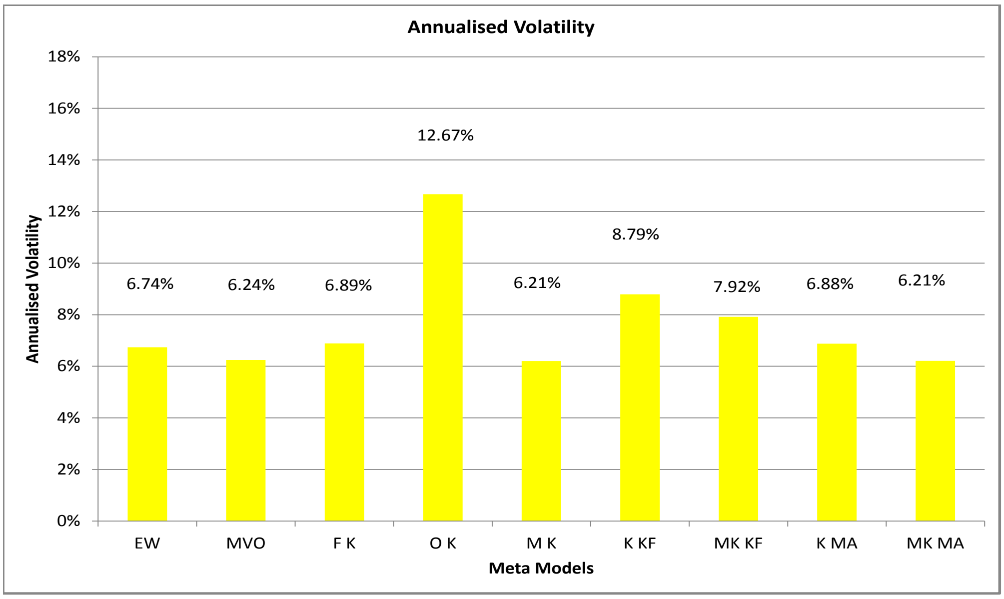


Figure D.2: Volatility of all the meta models during the in-sample period 2005-09. The Kelly models except Optimal Kelly exhibit lower volatility, even though its MVO which is supposed to capture the joint distirbution. The key reason for this is that the correlation between the QIM is low since they are working on different time horizons and capturing different aspect of data.

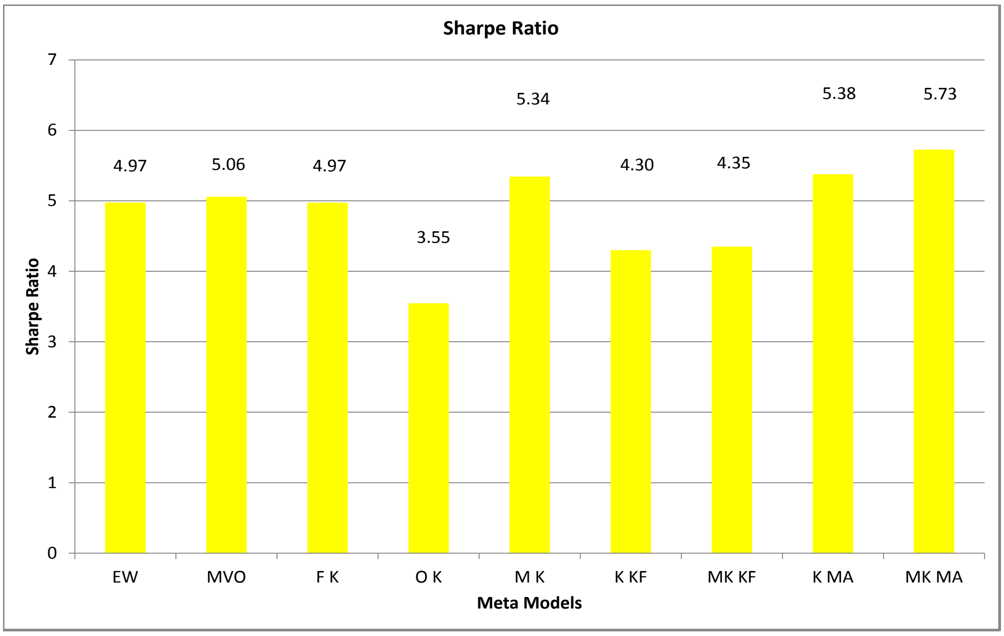


Figure D.3: Sharpe ratio of all the meta models during the in-sample period 2005-09. Kelly with Kalman Filter Moving Average and Median Kelly with Kalman Filter Moving Average arethe best performers the only models with a Sharpe ratio above 5, beating all benchmarks by some margin.

TABLE D.1: Comparison of Meta Model in the In-sample period

In- Sample	EW EW	MVO MVO	Fractional Kelly	Optimal Kelly	Median Kelly	K KF K KF	MK KF MK KF	K MA K MA	MK MA MK MA
Ann. Return	36.52%	34.57%	37.24%	47.92%	36.16%	40.80%	37.43%	39.99%	38.56%
Ann. Volatility	6.74%	6.24%	6.89%	12.67%	6.21%	8.79%	7.92%	6.88%	6.21%
Sharpe Ratio	4.97	5.06	4.97	3.55	5.34	4.30	4.35	5.38	5.73
Prob of gain	0.96	0.96	0.98	0.84	0.96	0.96	0.98	0.98	0.98
Prob. Of Loss	4.00%	4.00%	2.00%	16.00%	4.00%	4.00%	2.00%	0.02	0.02
Avg. win	2.74%	2.63%	2.73%	4.28%	2.72%	3.05%	2.78%	2.90%	2.82%
Avg. loss	-0.16%	-0.39%	-0.23%	-1.74%	-0.09%	-1.00%	-1.89%	-0.19%	-0.28%
Gain / Loss ratio	17.55	6.65	11.83	2.46	31.92	3.06	1.47	15.41	9.89
Median	2.36%	2.02%	2.29%	3.03%	2.17%	2.46%	2.47%	2.47%	2.58%
Average gain	2.63%	2.51%	2.67%	3.32%	2.61%	2.89%	2.69%	2.84%	2.75%

Key performance statistics for the Meta Models, including our three benchmarks. We can see the Kelly models perform well.

Appendix E

Out-of-sample meta model Performance

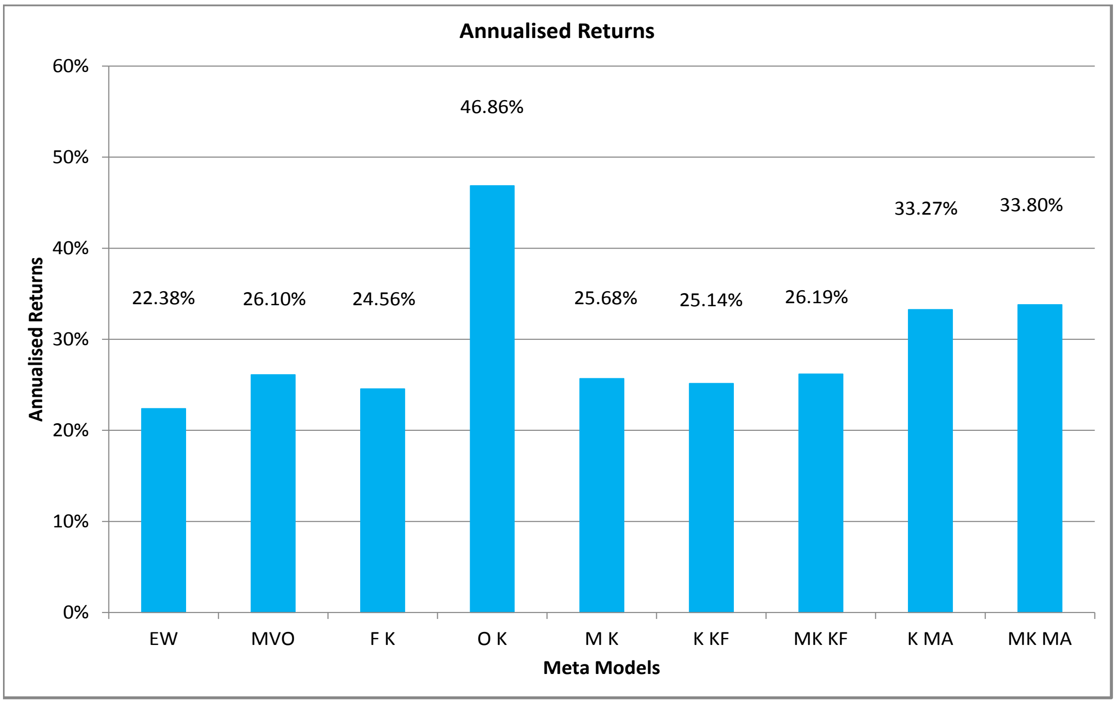


FIGURE E.1: Out-of-sample Annulaised returns for all the meta models. Just as we saw in the simulations and the In-sample period that Optimal Kelly has the best returns followed closely by Kelly with Kalman Filter Moving Average and Median Kelly with Kalman Filter Moving Average, both of which beat the rest opf the models by more that a full percentage point.

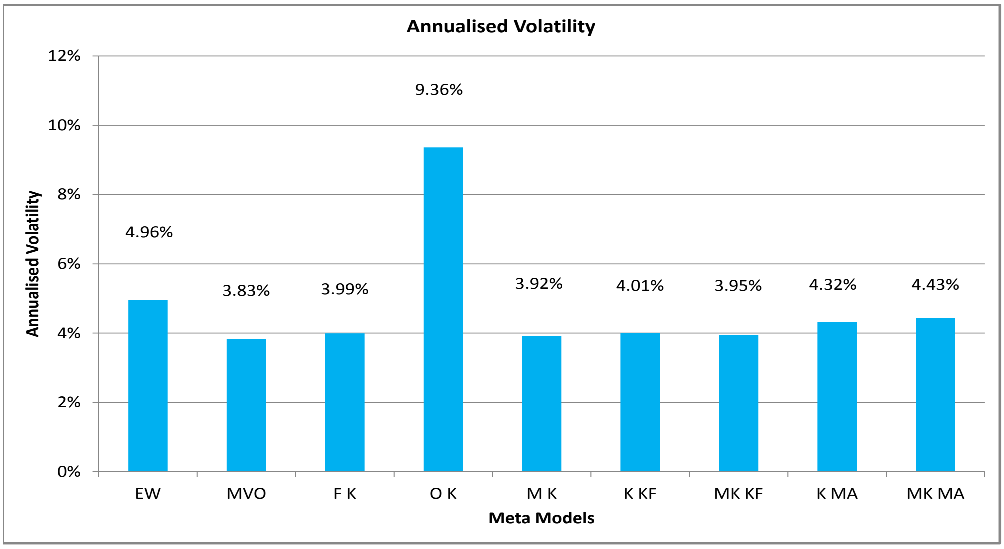


FIGURE E.2: Out-of-sample Annualised Volatility. As expected MVO has the lowest volatility very closely followed by Kelly with Kalman Filter Moving Average and Median Kelly with Kalman Filter Moving Average.

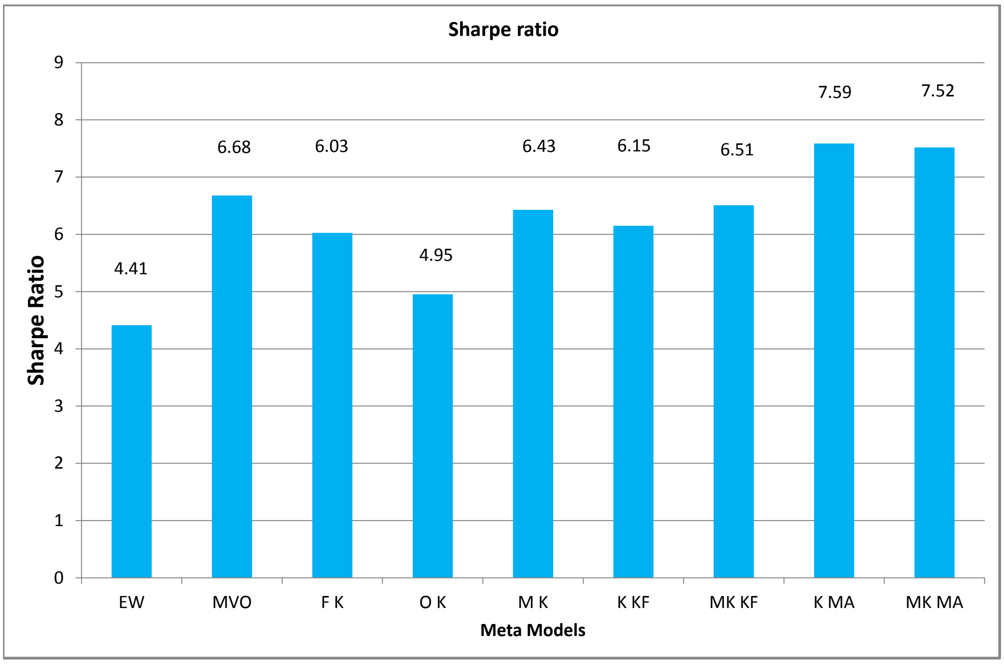


FIGURE E.3: Out-of-sample Sharpe Ratio. Kelly with Kalman Filter Moving Average and Median Kelly with Kalman Filter Moving Average have the the highest Sharpe ratio, in line with Simulations as well as the in-sample tests, showing that Kelly with Kalman Filter Moving Average and Median Kelly with Kalman Filter Moving Average are better at generating risk adjusted returns.

Table E.1: Out-of-sample Meta model comparison

Out-of-Sample	EW EW	MVO MVO	Fractional Kelly	Optimal Kelly	Median Kelly	K KF K KF	MK KF MK KF	K MA K MA	MK MA MK MA
Ann. Return	22.38%	26.10%	24.56%	46.86%	25.68%	25.14%	26.19%	33.27%	33.80%
Ann. Volatility	4.96%	3.83%	3.99%	9.36%	3.92%	4.01%	3.95%	4.32%	4.43%
Sharpe ratio	4.41	6.68	6.03	4.95	6.43	6.15	6.51	7.59	7.52
Prob of gain	0.92	0.94	0.94	0.92	0.94	0.94	0.94	0.97	0.97
Prob. Of Loss	0.08	0.06	0.06	0.08	0.06	0.06	5.56%	0.03	0.03
Avg. win	1.94%	2.09%	1.96%	3.84%	2.04%	2.00%	2.08%	2.49%	2.53%
Avg. loss	-0.97%	-0.47%	-0.13%	-3.16%	-0.14%	-0.10%	-0.11%	-0.03%	-0.16%
Gain / Loss ratio	2.00	4.50	15.00	1.21	14.43	20.44	18.11	95.84	15.54
Median	1.88%	1.95%	1.78%	3.03%	1.98%	1.81%	1.97%	2.46%	2.55%
Average gain	1.70%	1.95%	1.85%	3.25%	1.92%	1.89%	1.96%	2.42%	2.46%

Key performance numbers in the out-of-sample data set for the Meta Model, including our three benchmarks.

Appendix F

Meta model full sample

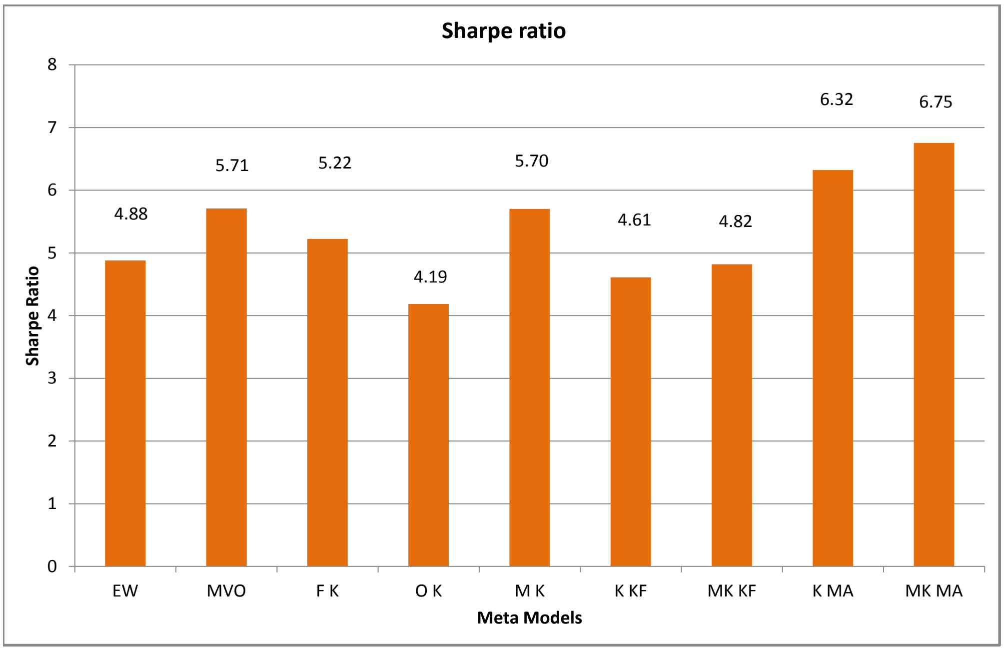


FIGURE F.1: In this chart we show full sample Sharpe Ratio. Kelly with Moving Average and Median Kelly with Moving Average have the the highest Sharpe ratio, in line with simulations as well as the in-sample tests, showing that Kelly with Moving Average and Median Kelly with Moving Average are better at generating risk adjustred returns.

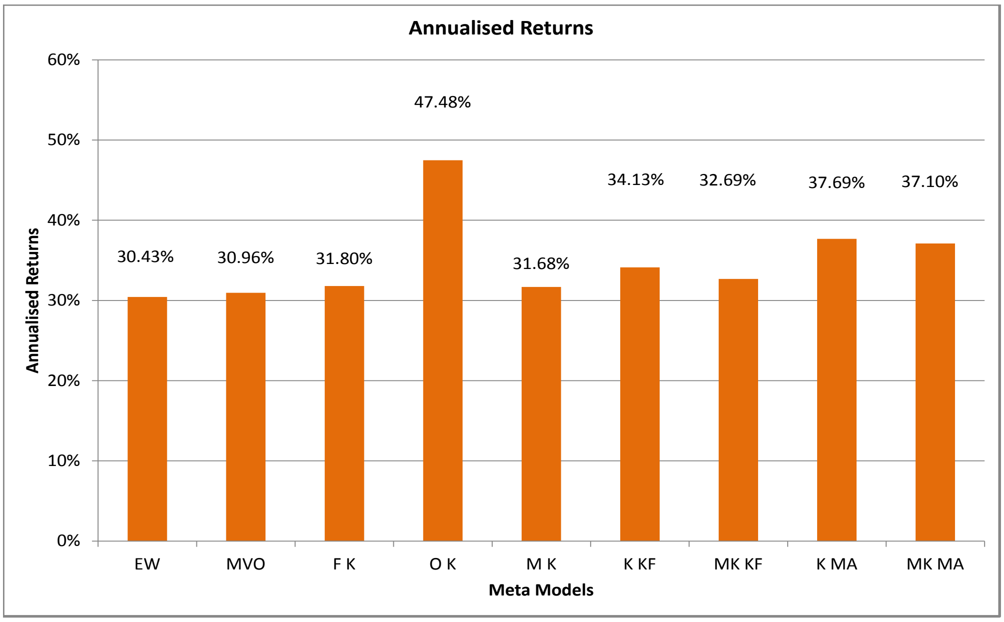


FIGURE F.2: In this chart we show full sample annualised returns. Optimal Kelly has the highed returns followed by Kelly with Moving Average and Median Kelly with Moving Average. MVO records the second lowest returns just ahead of equally weighted.

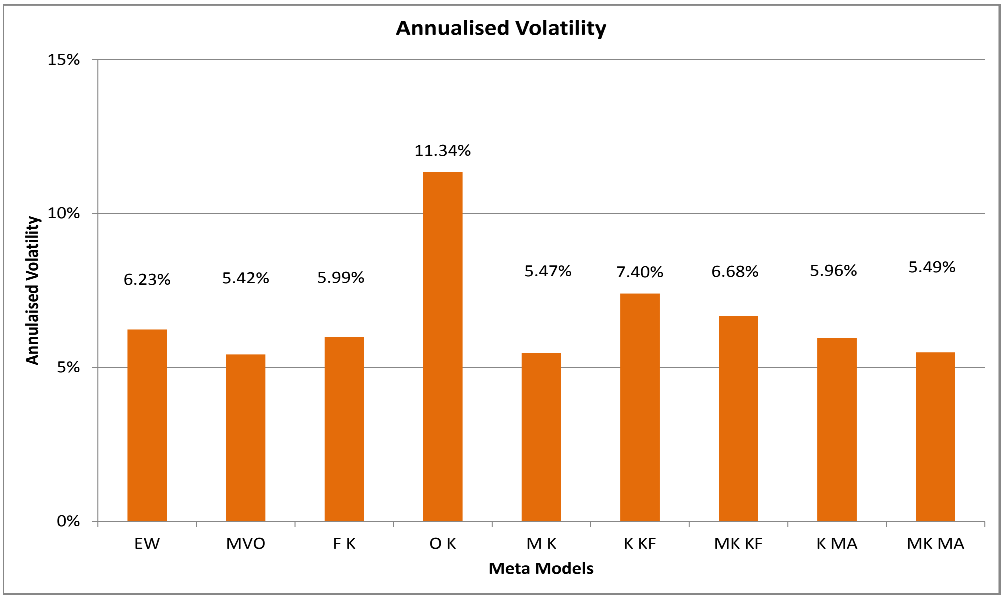


FIGURE F.3: In this chart we show full sample annualised volatiltiy. As expected MVO has the lowest volatility, followed by Median Kelly and Median Kelly with Moving Average. Optimal Kelly has the highest volatility

TABLE F.1: Full sample Meta model comparison

Full Sample	EW EW	MVO MVO	Fractional Kelly	Optimal Kelly	Median Kelly	K KF K KF	MK KF K2 MK	K MA K MA	MK MA MK MA
Ann. Return	30.43%	30.96%	31.80%	47.48%	31.68%	34.13%	32.69%	37.69%	37.10%
Ann. Volatility	6.23%	5.42%	5.99%	11.34%	5.47%	7.40%	6.68%	5.96%	5.49%
Sharpe ratio	4.88	5.71	5.22	4.19	5.70	4.61	4.82	6.32	6.75
Prob of gain	0.94	0.95	0.97	0.87	0.95	0.95	0.96	0.98	0.99
Prob. Of Loss	0.06	0.05	0.03	0.13	0.05	0.05	0.04	0.02	0.01
Avg. win	2.42%	2.41%	2.42%	4.09%	2.44%	2.63%	2.50%	2.77%	2.70%
Avg. loss	-0.64%	-0.43%	-0.16%	-2.13%	-0.11%	-0.55%	-0.71%	-0.11%	-0.16%
Gain / Loss ratio	3.75	5.59	14.72	1.92	21.51	4.79	3.54	25.83	16.57
Median	2.12%	2.02%	2.15%	3.03%	2.10%	2.10%	2.24%	2.47%	2.58%
Average gain	2.24%	2.27%	2.33%	3.29%	2.32%	2.48%	2.38%	2.70%	2.66%

Key performance numbers in the full sample data set for the Meta Model, including our three benchmarks.

Appendix G

Out-of-sample QIM Performance

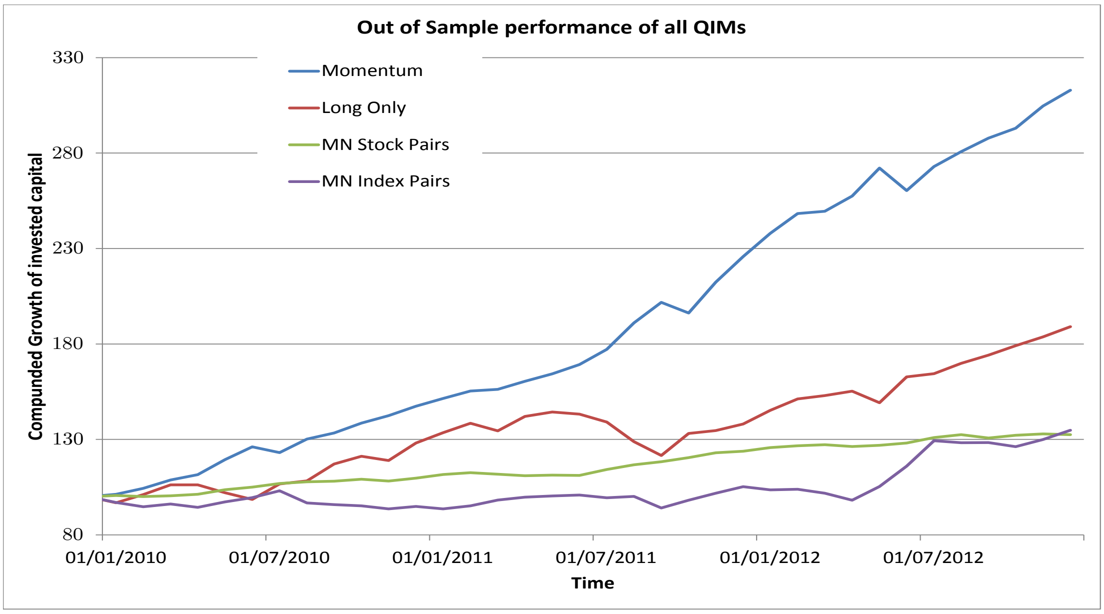


FIGURE G.1: Performance of the four quantitative investment models in the out-of-sample period. The Momentum model is the best performer.

Appendix H

Performance of QIM vs BAH of Investment Universe

Performance of the four QIMs when compared to buy and hold equally weighted return.

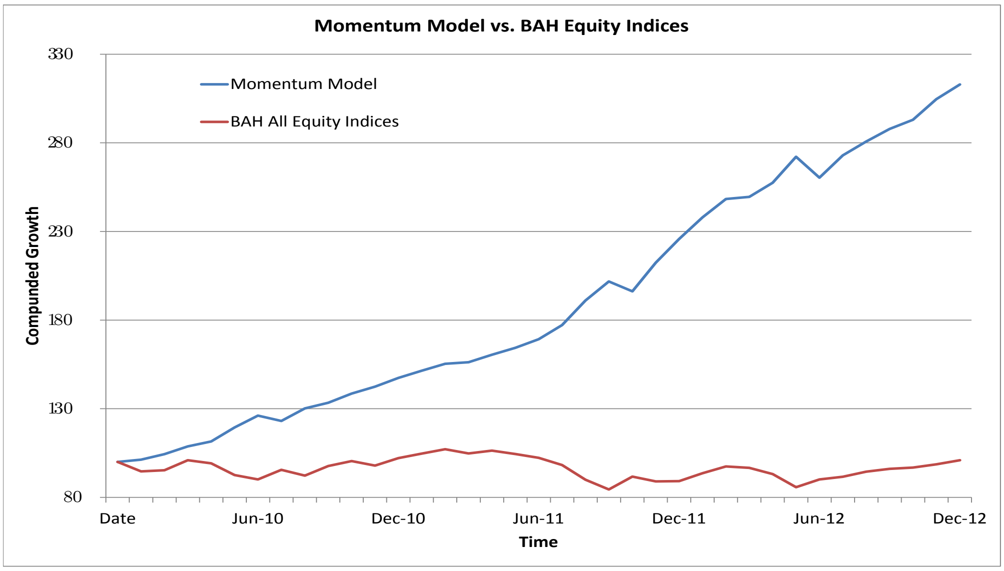


FIGURE H.1: Momentum model vs. BAH all Equity Indices

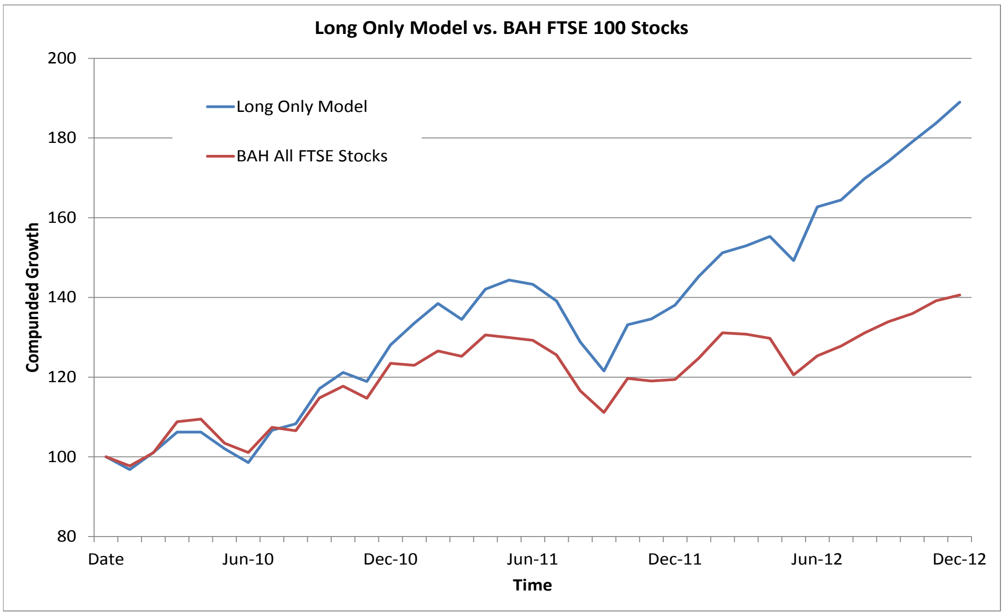


Figure H.2: Long Only model vs. BAH all UK FTSE 100 stocks

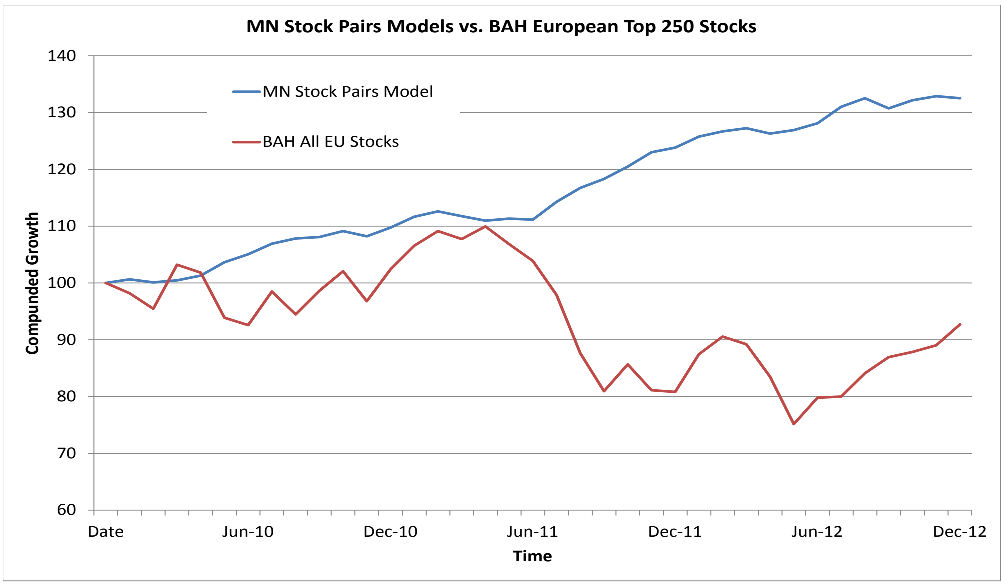


Figure H.3: MN Stock model vs. BAH all 250 European stocks

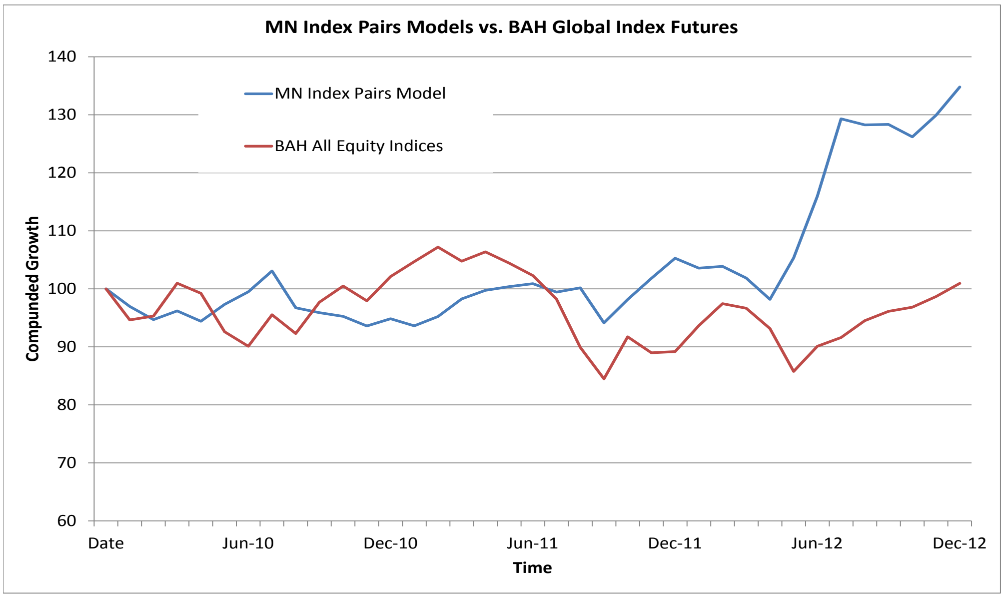


FIGURE H.4: MN Index model vs. BAH all Equity Indices

Appendix I

In-sample Distributions of QIMs

The bin size is calculated using Freedman - Diaconis method.

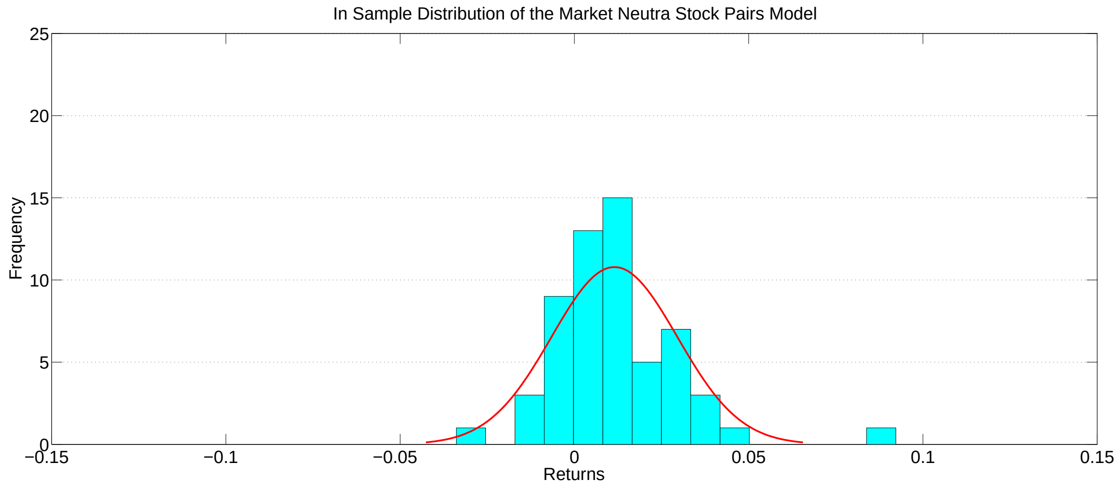


FIGURE I.1: In sample distribution of the Market Neutral Stock Pairs Model, number of bins = 15, bin size = 0.00840, mean = 0.01148, variance = 0.00032, skew = 1.51546, kurtosis = 6.43802, Jarque-Bera statistics is 103.34507.

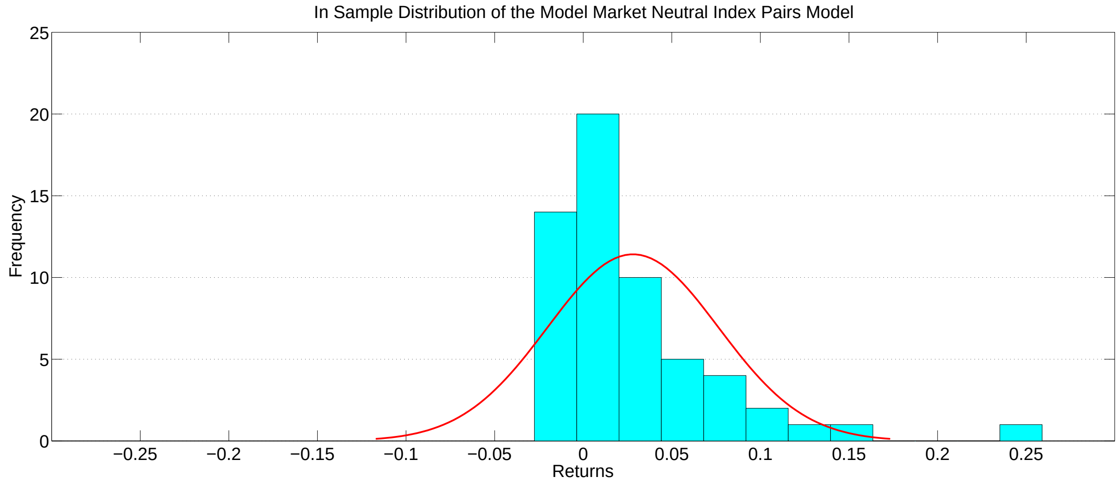


FIGURE I.2: In sample distribution of the Market Neutral Index Pairs Model, number of bins = 12, bin size = 0.02387, mean = 0.02808, variance = 0.00234, skew = 2.36665, kurtosis = 10.71947, Jarque-Bera statistics is 198.15330.

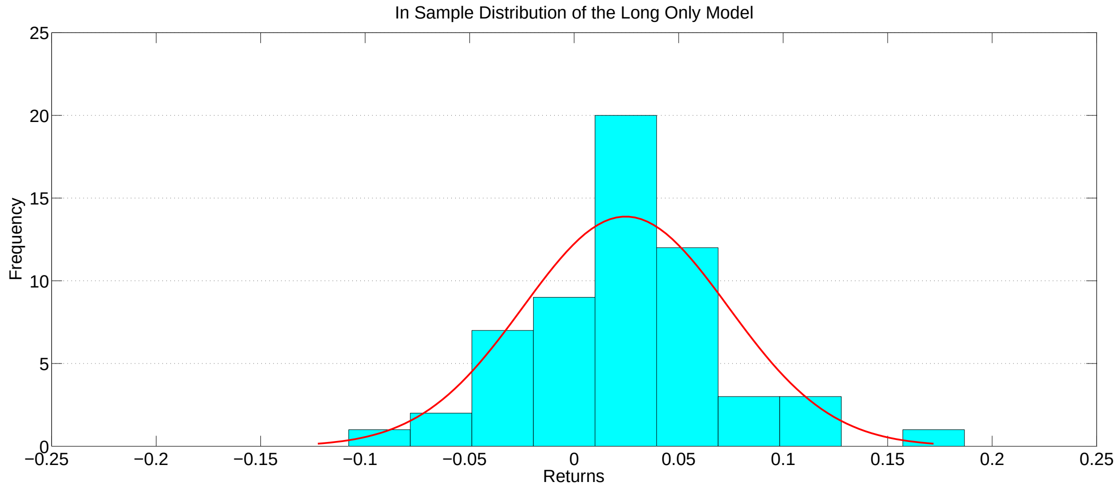


FIGURE I.3: In sample distribution of the Long Only Model, number of bins = 10, bin size = 0.02945, mean = 0.02464, variance = 0.00241, skew = 0.36463, kurtosis = 1.74318, Jarque-Bera statistics is 6.68507.

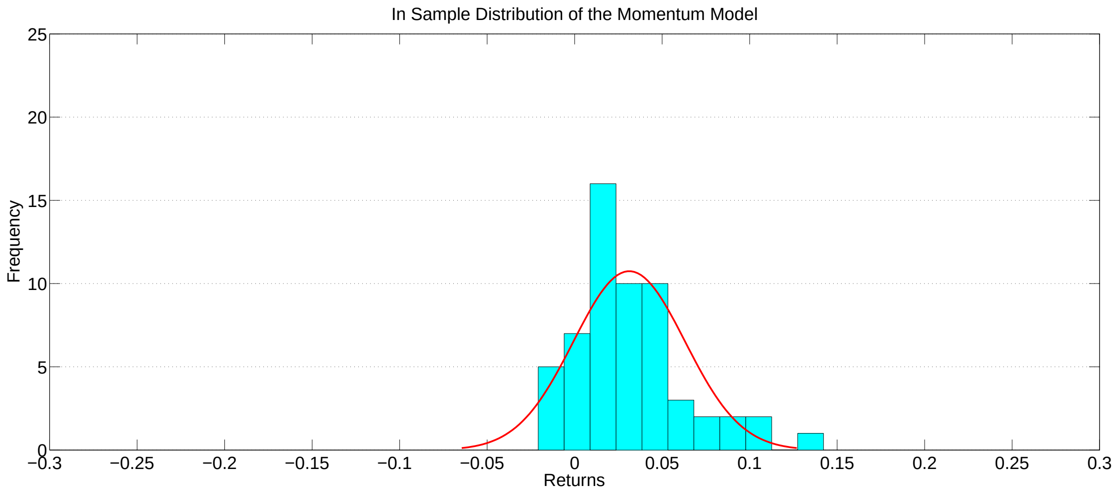


FIGURE I.4: In sample distribution of the Momentum Model, number of bins = 11, bin size = 0.01482, mean = 0.03122, variance = 0.00101, skew = 1.11545, kurtosis = 1.89122, Jarque-Bera statistics is 18.45137.

Appendix J

Out-of-sample Distributions of QIMs

The bin size is calculated using Freedman - Diaconis method.

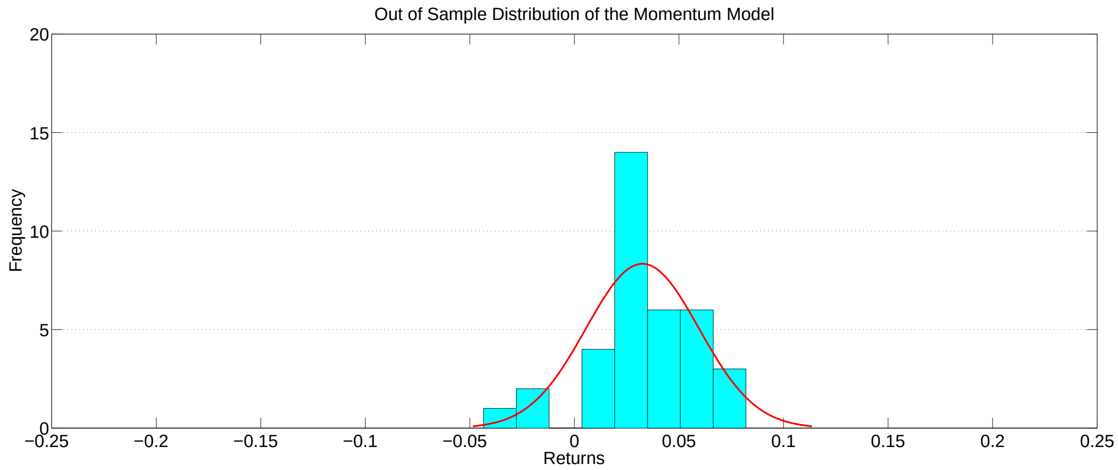


FIGURE J.1: Out of sample distribution of the Momentum model, number of bins = 8, bin size = 0.01569 mean = 0.03254, variance = 0.00072, skew = −0.75216, kurtosis = 1.36680, Jarque-Bera statistic = 4.96139.

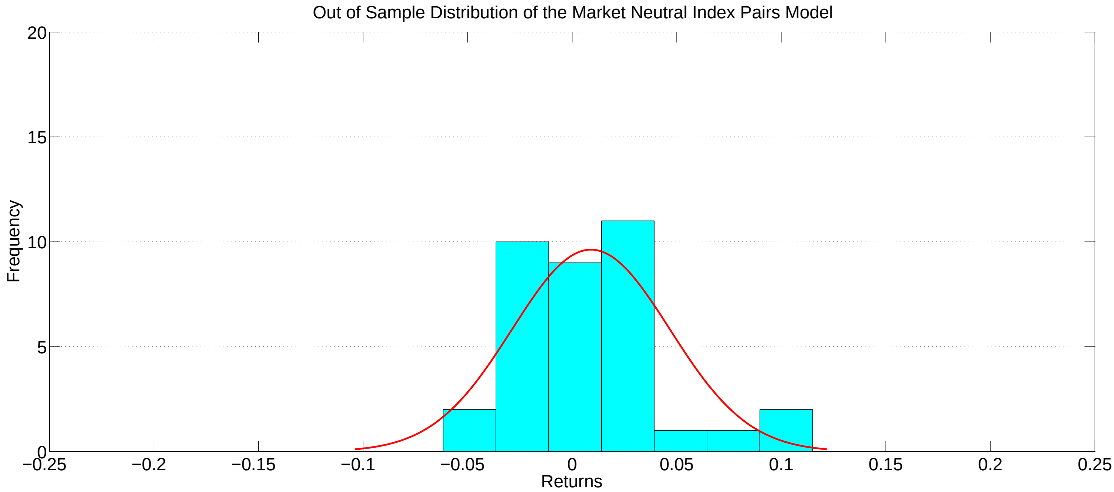


FIGURE J.2: Out of sample distribution of the MN Index Pairs model, number of bins = 7, bin size = 0.02524 mean = 0.00899, variance = 0.00141, skew = −0.74491, kurtosis = 1.40246, Jarque-Bera statistic = 4.99247.

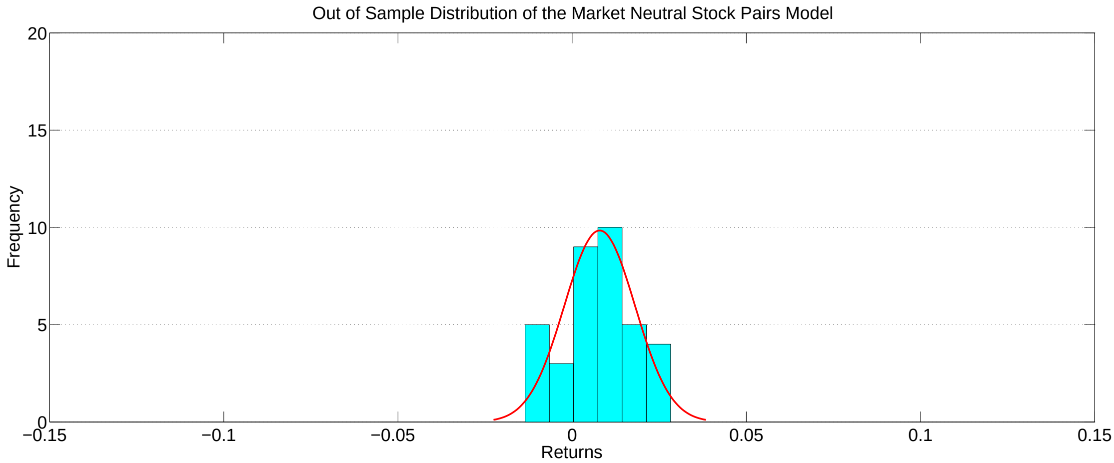


FIGURE J.3: Out of sample distribution of the MN Stock Pairs model, number of bins = 7, bin size = 0.00696 mean = 0.00899, variance = 0.00010, skew = −0.15275, kurtosis = −0.49803, Jarque-Bera statistic = 0.66866.

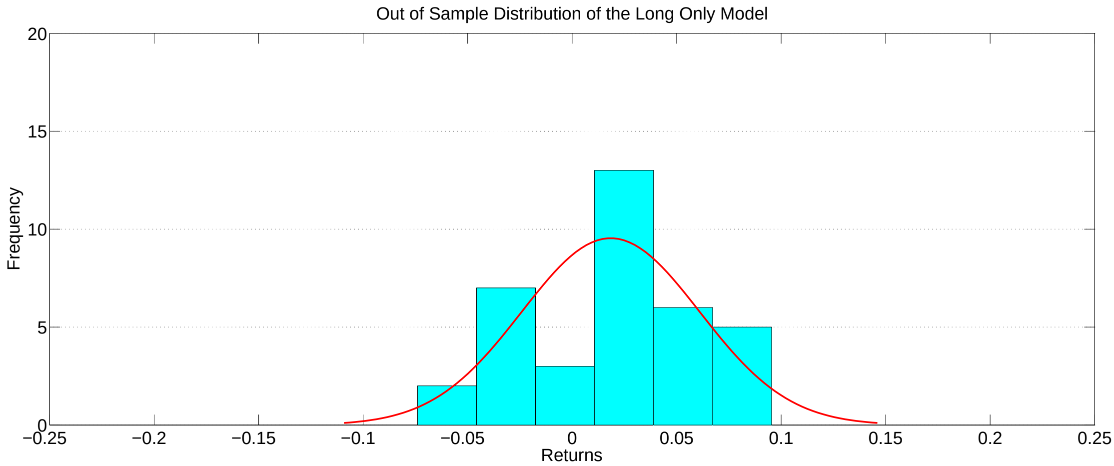


FIGURE J.4: Out of sample distribution of the Long Only model, number of bins = 6, bin size = 0.02824 mean = 0.01844, variance = 0.00180, skew = −0.18980, kurtosis = −0.46208, Jarque-Bera statistic = 0.69080.

Appendix K

Distributions of Meta Models

The distributions are calculated on the full data set. The bin size is calculated using Freedman - Diaconis method.

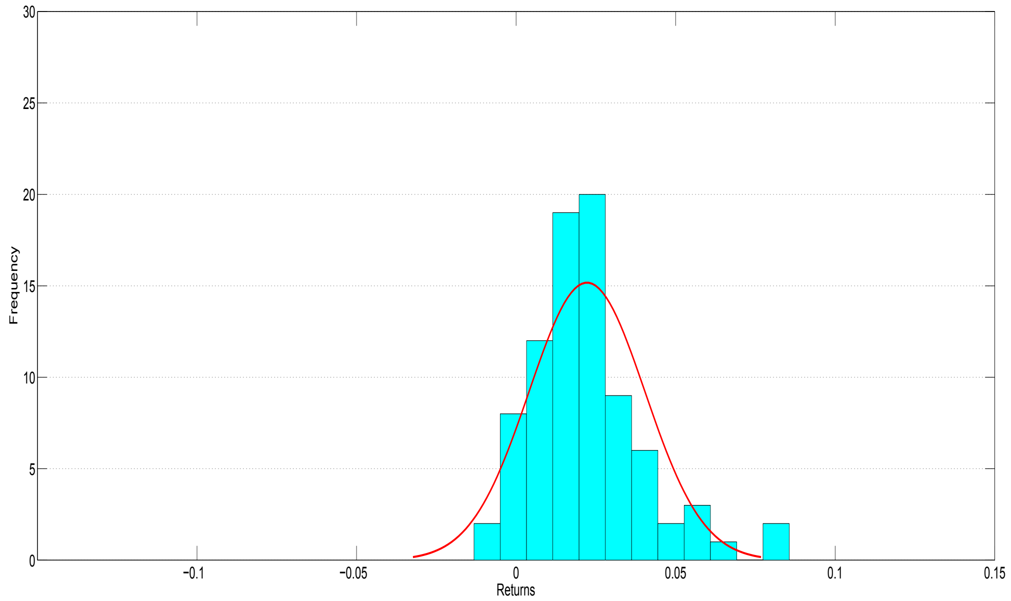


FIGURE K.1: Equally Weighted model distribution, number of bins = 12, bin size = 0.00823 mean = 0.02222, variance = 0.00033, skew = 1.01193, kurtosis = 4.67829.

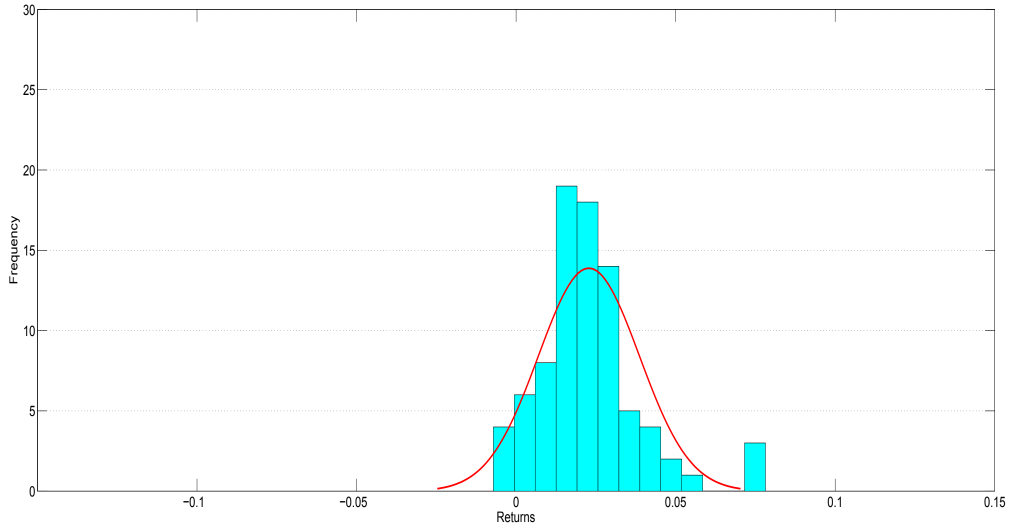


FIGURE K.2: MVO model distribution, number of bins = 13, mean = 0.02281, variance = 0.00025, skew = 1.22134, kurtosis = 5.63224.

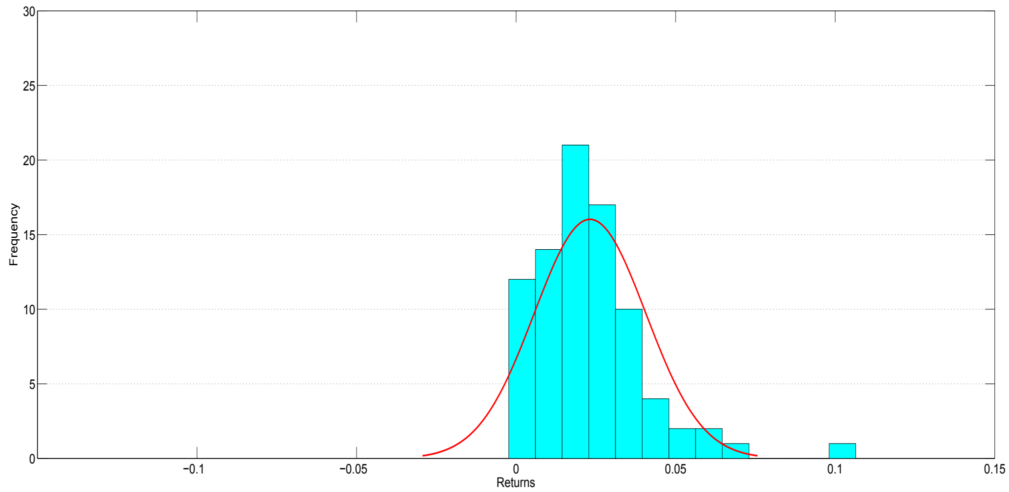


FIGURE K.3: Fractional Kelly model distributions, mean = 0.02317 = variance = 0.00030, skew = kurtosis = 8.03192

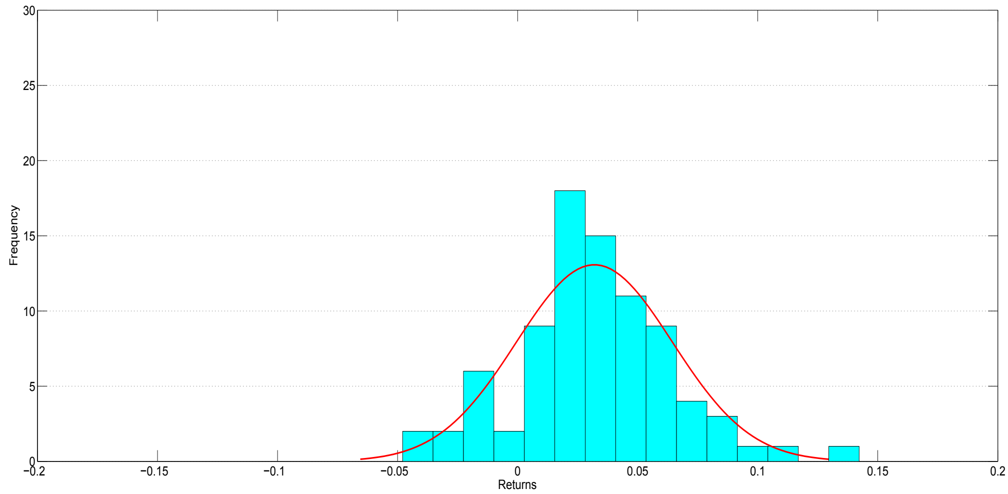


FIGURE K.4: Optimal Kelly model distribution, number of bins = 15, bin size = 0.01267, mean = 0.03198, variance = 0.00105, skew = 0.32971, kurtosis = 4.13955.

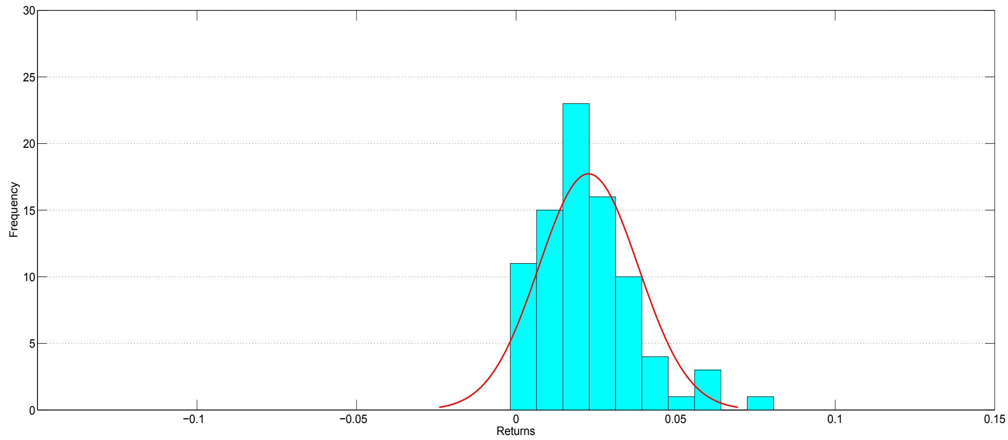


FIGURE K.5: Median Kelly model distributions, number of bins = 10, bin size = 0.00825, variance = 0.00024, skew = 1.06068, kurtosis = 4.73307.

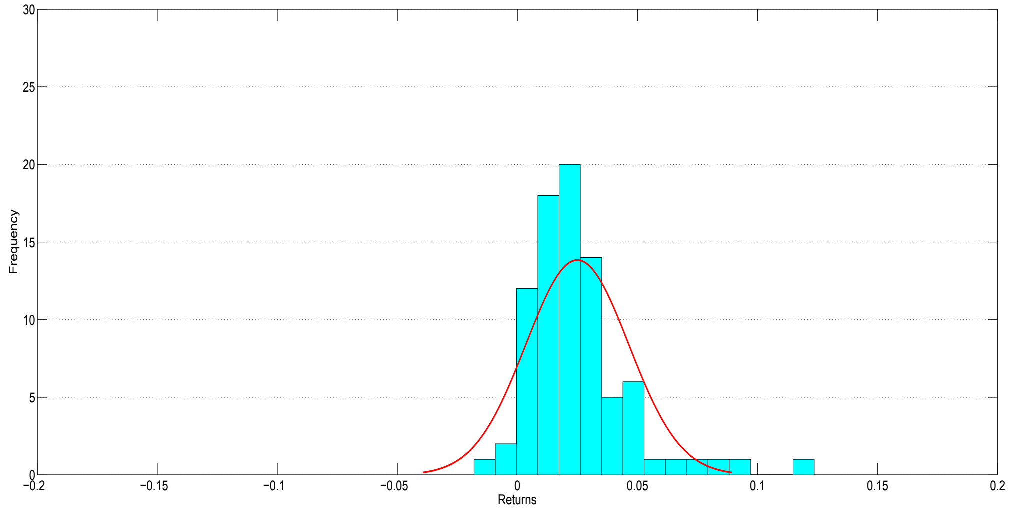


FIGURE K.6: Kelly with Kalman Filter model distributions, number of bins =16, bin size = 0.00885, variance = 0.00045, kurtosis = 8.15766.

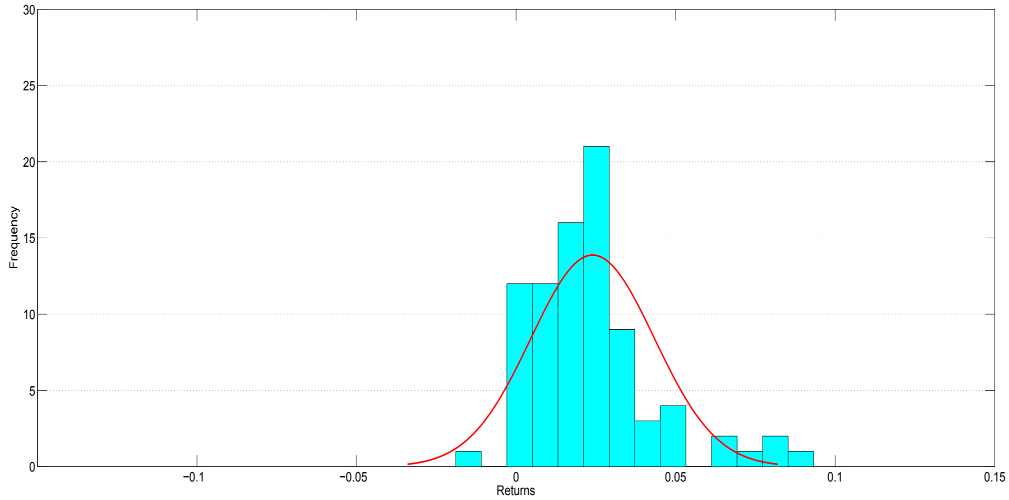


FIGURE K.7: Median Kelly with Kalman Filter model distributions, number of bins = 14, bin size = 0.00801, mean = 0.02399, variance = 0.00037, skew = 1.22792, kurtosis = 5.25810.

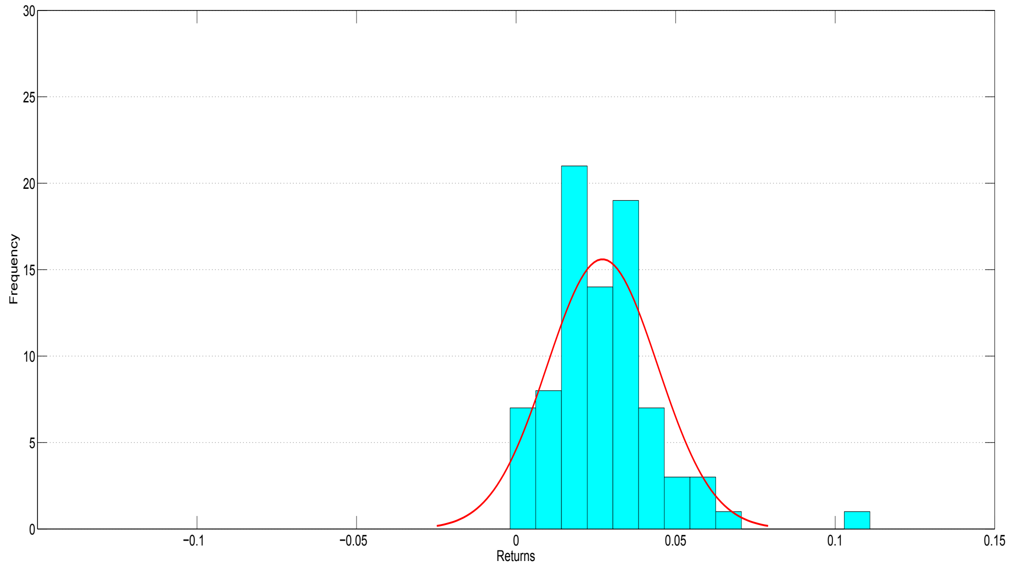


FIGURE K.8: Kelly with Moving Average model distributions, number of bins = 14, bin size - 0.00805, mean = 0.02706, variance = 0.00029 kurtosis = 8.34484.

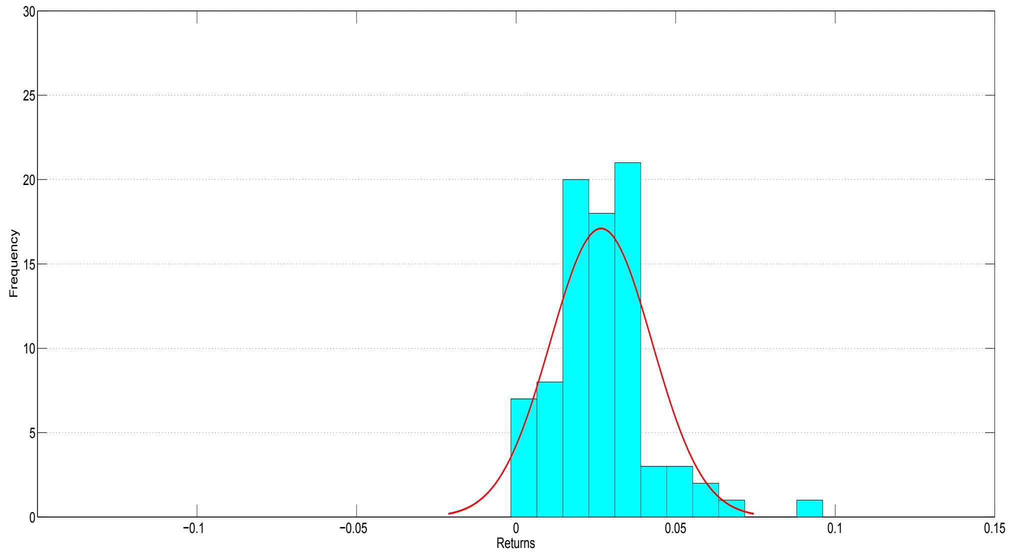


FIGURE K.9: Median Kelly with Moving Average model distributions, number of bins = 12, bin size = 0.00814, mean = 0.02664, variance = 0.00025, skew = 1.21625, kurtosis = 6.32661.

Appendix L

Kelly Weights for all Models

Here we present a graphical representation of weights assigned by all meta models for both the in-sample and out-of-sample period.

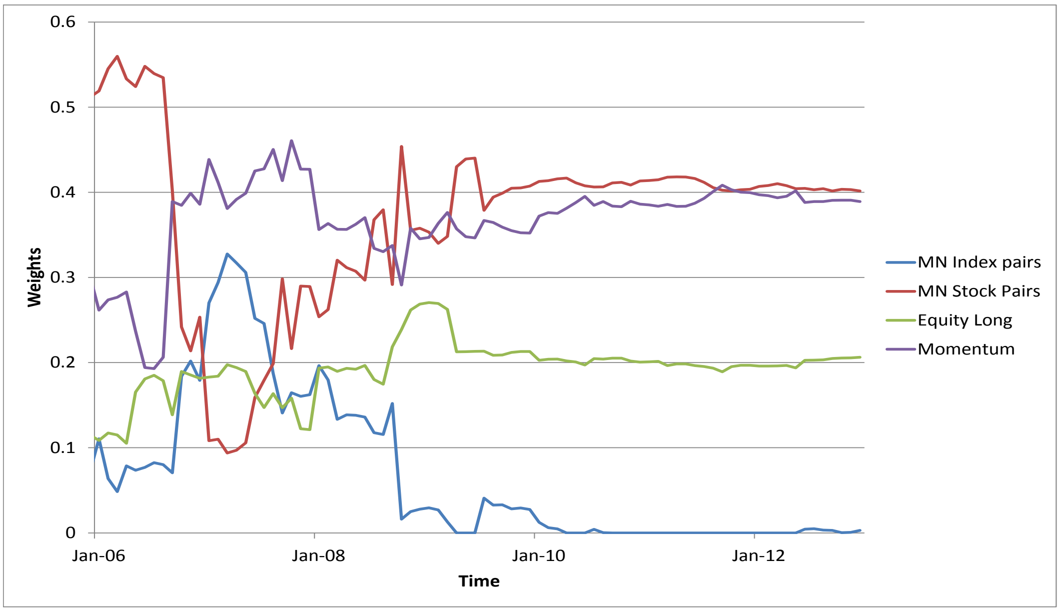


FIGURE L.1: Mean Variance framework weights.

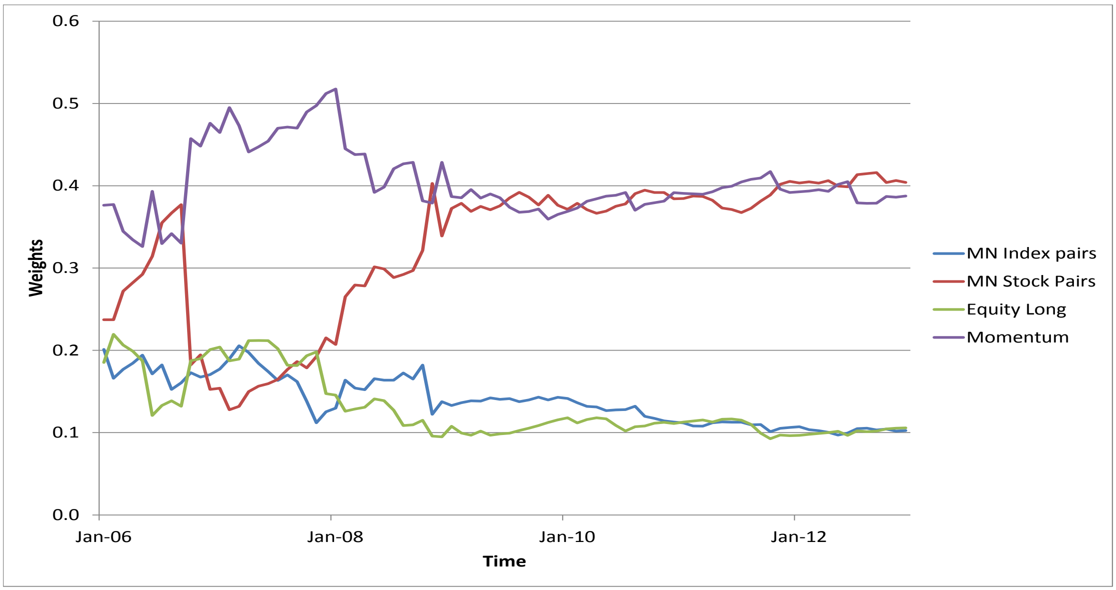


FIGURE L.2: Fractional Kelly weights.

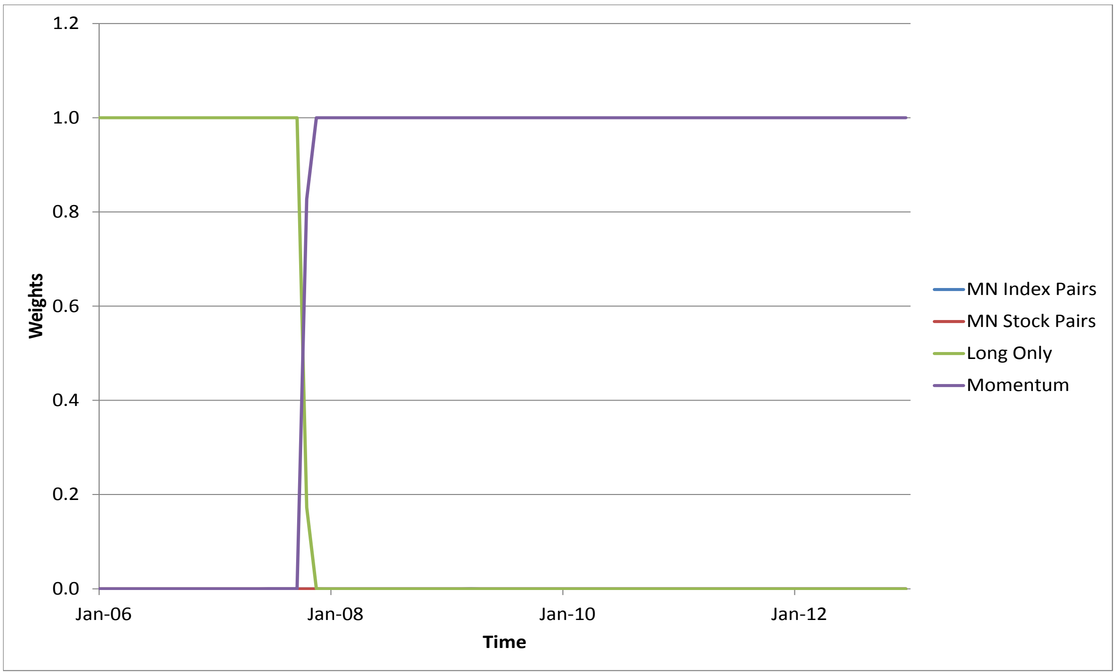


FIGURE L.3: Optimal Kelly weights.

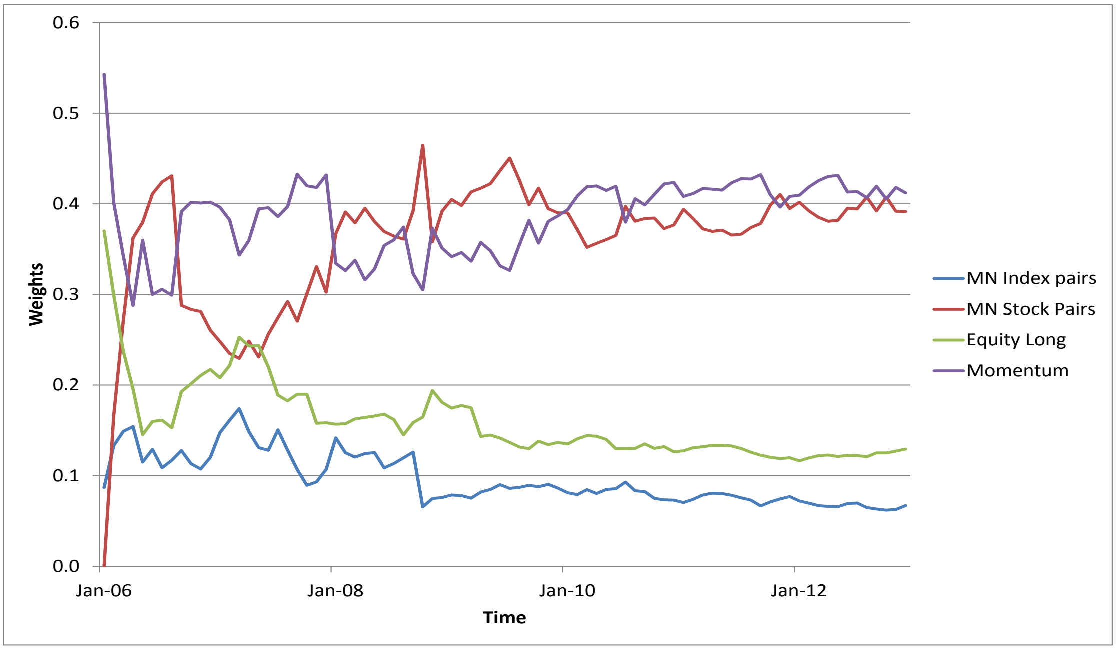


FIGURE L.4: Median Kelly weights.

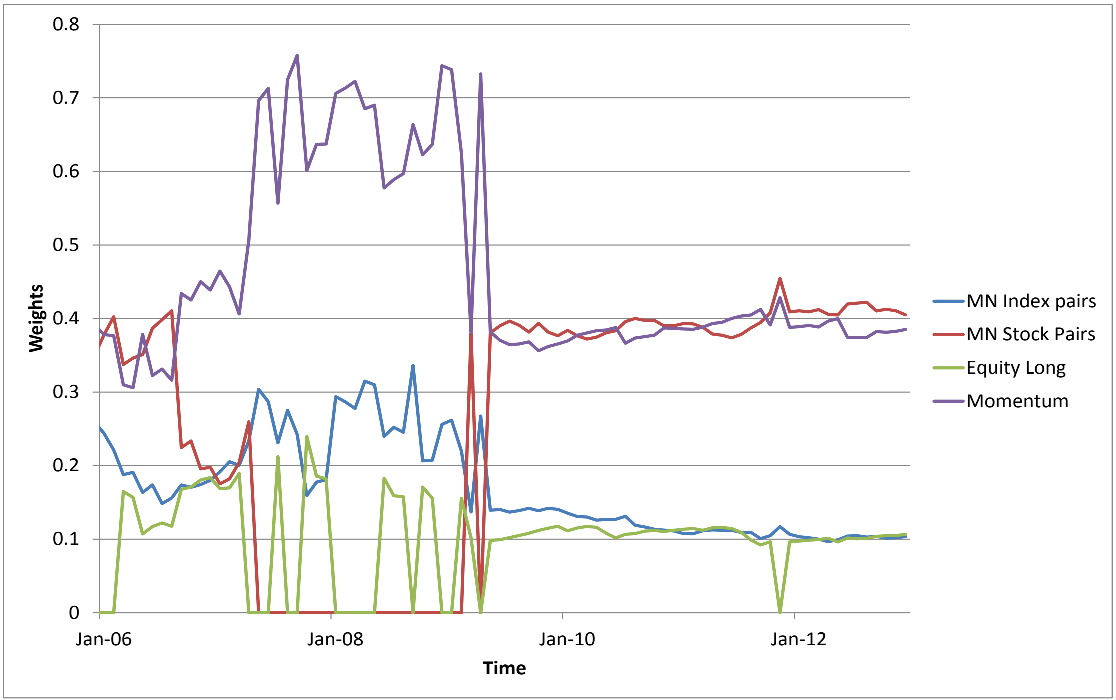


FIGURE L.5: Kelly with Kalman Filter weights.

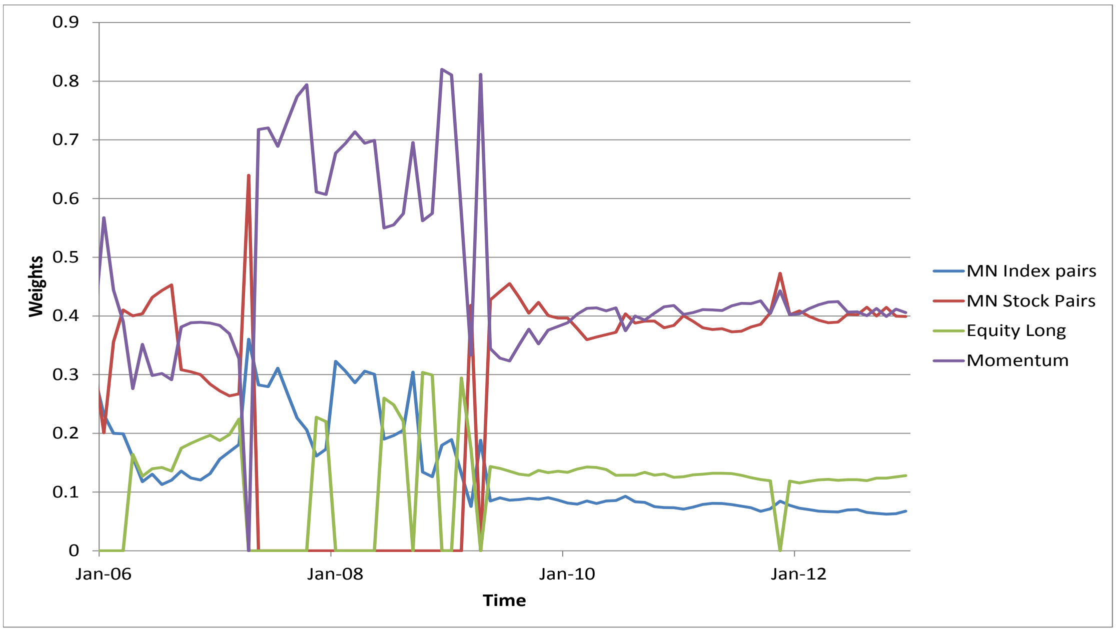


FIGURE L.6: Median Kelly with Kalman Filter weights.

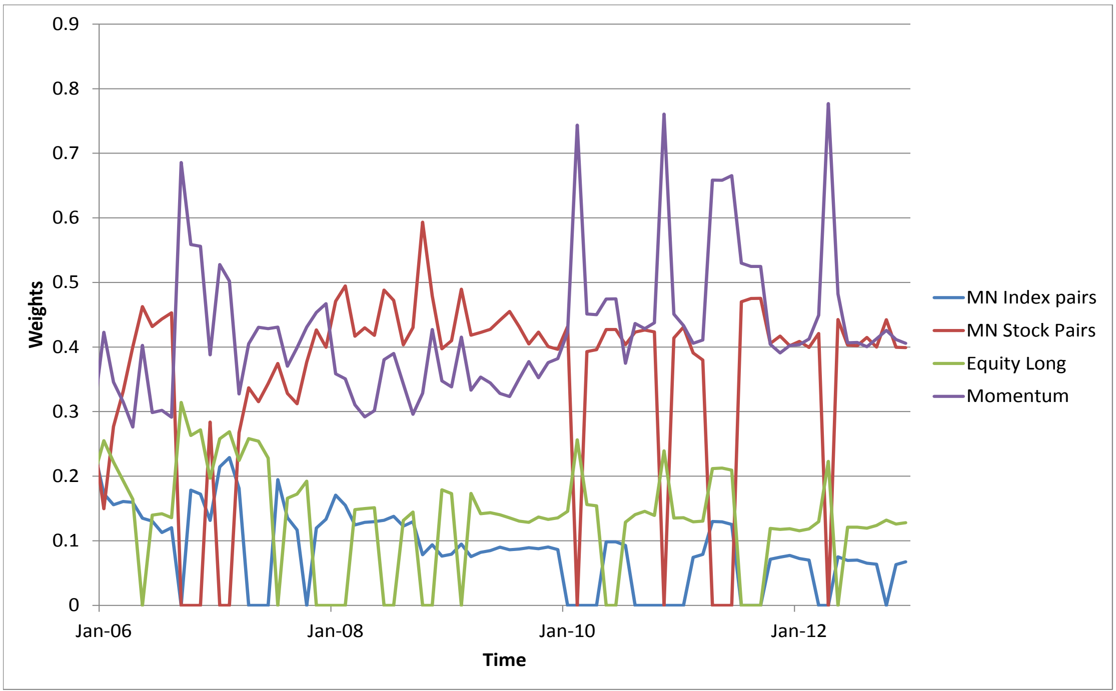


FIGURE L.7: Kelly with Moving Average weights.

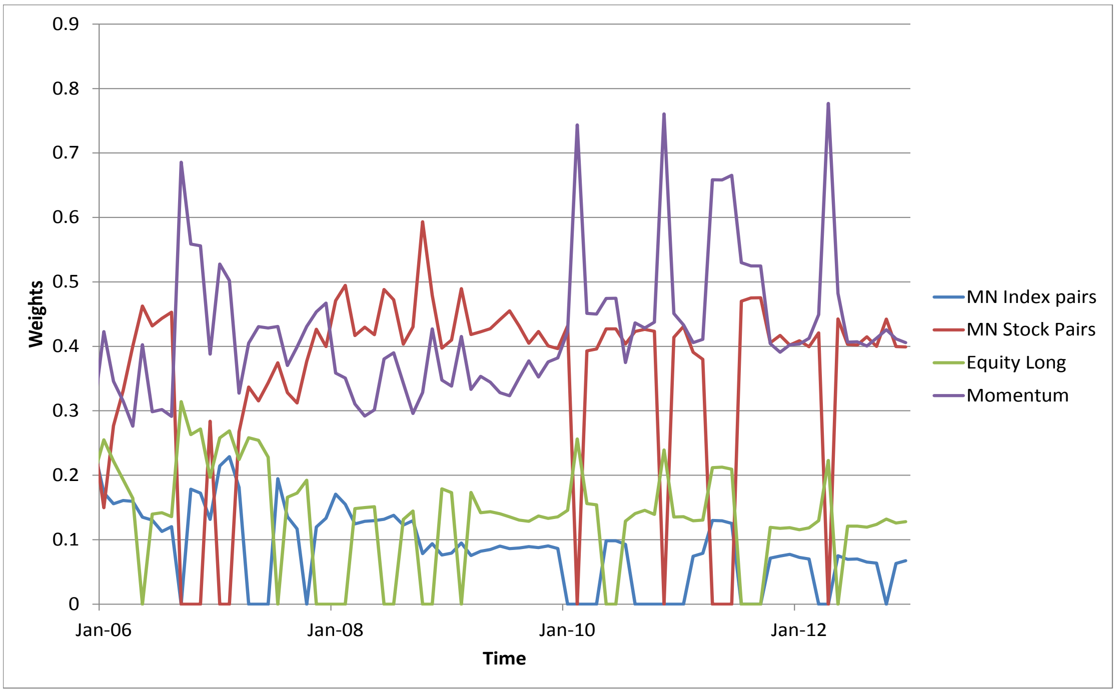


FIGURE L.8: Median Kelly with Moving Average weights.

Appendix M

Weights Maximising Gain in MVO & Optimal Kelly Weights

Table M.1: Optimal Kelly weights

Date	MN Index Pairs	MN Stock Pairs	Long Only	Momentum
Mar-05	0.99992	0.00003	0.00000	0.00005
Apr-05	0.00006	0.99993	0.00000	0.00001
May-05	0.00001	0.99997	0.00001	0.00001
Jun-05	0.00000	0.22242	0.77758	0.00000
Jul-05	0.00000	0.00000	0.99999	0.00000
Aug-05	0.00000	0.00000	0.99999	0.00000
Sep-05	0.00000	0.00000	1.00000	0.00000
Oct-05	0.00000	0.00000	0.99999	0.00000
Nov-05	0.00000	0.00000	0.99999	0.00000
Dec-05	0.00000	0.00000	1.00000	0.00000
Jan-06	0.00000	0.00000	1.00000	0.00000
Feb-06	0.00000	0.00000	1.00000	0.00000
Mar-06	0.00000	0.00000	1.00000	0.00000
Apr-06	0.00000	0.00000	1.00000	0.00000
May-06	0.00000	0.00000	1.00000	0.00000
Jun-06	0.00000	0.00000	1.00000	0.00000
Jul-06	0.00000	0.00000	1.00000	0.00000
Aug-06	0.00000	0.00000	1.00000	0.00000
Sep-06	0.00000	0.00000	1.00000	0.00000
Oct-06	0.00000	0.00000	1.00000	0.00000
Nov-06	0.00000	0.00000	1.00000	0.00000
Dec-06	0.00000	0.00000	1.00000	0.00000
Jan-07	0.00000	0.00000	0.99999	0.00001
Feb-07	0.00000	0.00000	0.99999	0.00001
Mar-07	0.00000	0.00000	0.99999	0.00001
Apr-07	0.00000	0.00000	0.99999	0.00001
May-07	0.00000	0.00000	0.99999	0.00001
Jun-07	0.00000	0.00000	0.99999	0.00001
Jul-07	0.00000	0.00000	0.99999	0.00001

Table M.2: Optimal Kelly weights

Aug-07	0.00000	0.00000	0.99996	0.00004
Sep-07	0.00000	0.00000	0.99994	0.00006
Oct-07	0.00000	0.00000	0.17260	0.82740
Nov-07	0.00000	0.00000	0.00001	0.99999
Dec-07	0.00000	0.00000	0.00001	0.99999
Jan-08	0.00000	0.00000	0.00000	1.00000
Feb-08	0.00000	0.00000	0.00000	1.00000
Mar-08	0.00000	0.00000	0.00000	1.00000
Apr-08	0.00000	0.00000	0.00000	1.00000
May-08	0.00000	0.00000	0.00000	1.00000
Jun-08	0.00000	0.00000	0.00000	1.00000
Jul-08	0.00000	0.00000	0.00000	1.00000
Aug-08	0.00000	0.00000	0.00000	1.00000
Sep-08	0.00000	0.00000	0.00000	1.00000
Oct-08	0.00000	0.00000	0.00000	1.00000
Nov-08	0.00000	0.00000	0.00000	1.00000
Dec-08	0.00000	0.00000	0.00000	0.99999
Jan-09	0.00001	0.00000	0.00000	0.99999
Feb-09	0.00000	0.00000	0.00000	1.00000
Mar-09	0.00001	0.00000	0.00000	0.99999
Apr-09	0.00001	0.00000	0.00000	0.99999
May-09	0.00001	0.00000	0.00000	0.99999
Jun-09	0.00000	0.00000	0.00000	1.00000
Jul-09	0.00000	0.00000	0.00000	1.00000
Aug-09	0.00000	0.00000	0.00000	1.00000
Sep-09	0.00000	0.00000	0.00000	0.99999
Oct-09	0.00000	0.00000	0.00000	1.00000
Nov-09	0.00000	0.00000	0.00000	0.99999
Dec-09	0.00000	0.00000	0.00000	1.00000
Jan-10	0.00000	0.00000	0.00000	1.00000
Feb-10	0.00000	0.00000	0.00000	1.00000
Mar-10	0.00000	0.00000	0.00000	1.00000
Apr-10	0.00000	0.00000	0.00000	1.00000

Table M.3: Optimal Kelly weights

May-10	0.00000	0.00000	0.00000	1.00000
Jun-10	0.00000	0.00000	0.00000	1.00000
Jul-10	0.00000	0.00000	0.00000	1.00000
Aug-10	0.00000	0.00000	0.00000	1.00000
Sep-10	0.00000	0.00000	0.00000	1.00000
Oct-10	0.00000	0.00000	0.00000	1.00000
Nov-10	0.00000	0.00000	0.00000	1.00000
Dec-10	0.00000	0.00000	0.00000	1.00000
Jan-11	0.00000	0.00000	0.00000	1.00000
Feb-11	0.00000	0.00000	0.00000	1.00000
Mar-11	0.00000	0.00000	0.00000	1.00000
Apr-11	0.00000	0.00000	0.00000	1.00000
May-11	0.00000	0.00000	0.00000	1.00000
Jun-11	0.00000	0.00000	0.00000	1.00000
Jul-11	0.00000	0.00000	0.00000	1.00000
Aug-11	0.00000	0.00000	0.00000	1.00000
Sep-11	0.00000	0.00000	0.00000	1.00000
Oct-11	0.00000	0.00000	0.00000	1.00000
Nov-11	0.00000	0.00000	0.00000	1.00000
Dec-11	0.00000	0.00000	0.00000	1.00000
Jan-12	0.00000	0.00000	0.00000	1.00000
Feb-12	0.00000	0.00000	0.00000	1.00000
Mar-12	0.00000	0.00000	0.00000	1.00000
Apr-12	0.00000	0.00000	0.00000	1.00000
May-12	0.00000	0.00000	0.00000	1.00000
Jun-12	0.00000	0.00000	0.00000	1.00000
Jul-12	0.00000	0.00000	0.00000	1.00000
Aug-12	0.00000	0.00000	0.00000	1.00000
Sep-12	0.00000	0.00000	0.00000	1.00000
Oct-12	0.00000	0.00000	0.00000	1.00000
Nov-12	0.00000	0.00000	0.00000	1.00000
Dec-12	0.00000	0.00000	0.00000	1.00000

Table M.4: MVO weights maximising gain

Date	MN Index Pairs	MN Stock Pairs	Long Only	Momentum
Apr-05	0.00000	1.00000	0.00000	0.00000
May-05	0.00000	1.00000	0.00000	0.00000
Jun-05	0.00000	0.00000	1.00000	0.00000
Jul-05	0.00000	0.00000	1.00000	0.00000
Aug-05	0.00000	0.00000	1.00000	0.00000
Sep-05	0.00000	0.00000	1.00000	0.00000
Oct-05	0.00000	0.00000	1.00000	0.00000
Nov-05	0.00000	0.00000	1.00000	0.00000
Dec-05	0.00000	0.00000	1.00000	0.00000
Jan-06	0.00000	0.00000	1.00000	0.00000
Feb-06	0.00000	0.00000	1.00000	0.00000
Mar-06	0.00000	0.00000	1.00000	0.00000
Apr-06	0.00000	0.00000	1.00000	0.00000
May-06	0.00000	0.00000	1.00000	0.00000
Jun-06	0.00000	0.00000	1.00000	0.00000
Jul-06	0.00000	0.00000	1.00000	0.00000
Aug-06	0.00000	0.00000	1.00000	0.00000
Sep-06	0.00000	0.00000	1.00000	0.00000
Oct-06	0.00000	0.00000	1.00000	0.00000
Nov-06	0.00000	0.00000	1.00000	0.00000
Dec-06	0.00000	0.00000	1.00000	0.00000
Jan-07	0.00000	0.00000	1.00000	0.00000
Feb-07	0.00000	0.00000	1.00000	0.00000
Mar-07	0.00000	0.00000	1.00000	0.00000
Apr-07	0.00000	0.00000	1.00000	0.00000
May-07	0.00000	0.00000	1.00000	0.00000
Jun-07	0.00000	0.00000	1.00000	0.00000
Jul-07	0.00000	0.00000	1.00000	0.00000
Aug-07	0.00000	0.00000	1.00000	0.00000

Table M.5: MVO weights maximising gain

Sep-07	0.00000	0.00000	1.00000	0.00000
Oct-07	0.00000	0.00000	0.00000	1.00000
Nov-07	0.00000	0.00000	0.00000	1.00000
Dec-07	0.00000	0.00000	0.00000	1.00000
Jan-08	0.00000	0.00000	0.00000	1.00000
Feb-08	0.00000	0.00000	0.00000	1.00000
Mar-08	0.00000	0.00000	0.00000	1.00000
Apr-08	0.00000	0.00000	0.00000	1.00000
May-08	0.00000	0.00000	0.00000	1.00000
Jun-08	0.00000	0.00000	0.00000	1.00000
Jul-08	0.00000	0.00000	0.00000	1.00000
Aug-08	0.00000	0.00000	0.00000	1.00000
Sep-08	0.00000	0.00000	0.00000	1.00000
Oct-08	0.00000	0.00000	0.00000	1.00000
Nov-08	0.00000	0.00000	0.00000	1.00000
Dec-08	0.00000	0.00000	0.00000	1.00000
Jan-09	0.00000	0.00000	0.00000	1.00000
Feb-09	0.00000	0.00000	0.00000	1.00000
Mar-09	0.00000	0.00000	0.00000	1.00000
Apr-09	0.00000	0.00000	0.00000	1.00000
May-09	0.00000	0.00000	0.00000	1.00000
Jun-09	0.00000	0.00000	0.00000	1.00000
Jul-09	0.00000	0.00000	0.00000	1.00000
Aug-09	0.00000	0.00000	0.00000	1.00000
Sep-09	0.00000	0.00000	0.00000	1.00000
Oct-09	0.00000	0.00000	0.00000	1.00000
Nov-09	0.00000	0.00000	0.00000	1.00000
Dec-09	0.00000	0.00000	0.00000	1.00000
Jan-10	0.00000	0.00000	0.00000	1.00000
Feb-10	0.00000	0.00000	0.00000	1.00000
Mar-10	0.00000	0.00000	0.00000	1.00000

Table M.6: MVO weights maximising gain

Apr-10	0.00000	0.00000	0.00000	1.00000
May-10	0.00000	0.00000	0.00000	1.00000
Jun-10	0.00000	0.00000	0.00000	1.00000
Jul-10	0.00000	0.00000	0.00000	1.00000
Aug-10	0.00000	0.00000	0.00000	1.00000
Sep-10	0.00000	0.00000	0.00000	1.00000
Oct-10	0.00000	0.00000	0.00000	1.00000
Nov-10	0.00000	0.00000	0.00000	1.00000
Dec-10	0.00000	0.00000	0.00000	1.00000
Jan-11	0.00000	0.00000	0.00000	1.00000
Feb-11	0.00000	0.00000	0.00000	1.00000
Mar-11	0.00000	0.00000	0.00000	1.00000
Apr-11	0.00000	0.00000	0.00000	1.00000
May-11	0.00000	0.00000	0.00000	1.00000
Jun-11	0.00000	0.00000	0.00000	1.00000
Jul-11	0.00000	0.00000	0.00000	1.00000
Aug-11	0.00000	0.00000	0.00000	1.00000
Sep-11	0.00000	0.00000	0.00000	1.00000
Oct-11	0.00000	0.00000	0.00000	1.00000
Nov-11	0.00000	0.00000	0.00000	1.00000
Dec-11	0.00000	0.00000	0.00000	1.00000
Jan-12	0.00000	0.00000	0.00000	1.00000
Feb-12	0.00000	0.00000	0.00000	1.00000
Mar-12	0.00000	0.00000	0.00000	1.00000
Apr-12	0.00000	0.00000	0.00000	1.00000
May-12	0.00000	0.00000	0.00000	1.00000
Jun-12	0.00000	0.00000	0.00000	1.00000
Jul-12	0.00000	0.00000	0.00000	1.00000
Aug-12	0.00000	0.00000	0.00000	1.00000
Sep-12	0.00000	0.00000	0.00000	1.00000
Oct-12	0.00000	0.00000	0.00000	1.00000
Nov-12	0.00000	0.00000	0.00000	1.00000
Dec-12	0.00000	0.00000	0.00000	1.00000

Appendix N

FTSE 100 Stocks

List of stocks that make up the FTSE 100.

Table N.1: FTSE 100 stocks

Number	Ticker	Number	Ticker	Number	Ticker
1	AAL LN Equity	34	FRES LN Equity	67	RDSA LN Equity
2	ABF LN Equity	35	GFS LN Equity	68	RDSB LN Equity
3	ADM LN Equity	36	GKN LN Equity	69	REL LN Equity
4	ADN LN Equity	37	GLEN LN Equity	70	REX LN Equity
5	AGK LN Equity	38	GSK LN Equity	71	RIO LN Equity
6	AMEC LN Equity	39	HL/ LN Equity	72	RR/ LN Equity
7	ANTO LN Equity	40	HMSO LN Equity	73	RRS LN Equity
8	ARM LN Equity	41	HSBA LN Equity	74	RSA LN Equity
9	AV/ LN Equity	42	IAG LN Equity	75	RSL LN Equity
10	AZN LN Equity	43	IHG LN Equity	76	SAB LN Equity
11	BA/ LN Equity	44	IMI LN Equity	77	SBRY LN Equity
12	BAB LN Equity	45	IMT LN Equity	78	SDR LN Equity
13	BARC LN Equity	46	ITRK LN Equity	79	SGE LN Equity
14	BATS LN Equity	47	ITV LN Equity	80	SHP LN Equity
15	BG/ LN Equity	48	JMAT LN Equity	81	SL/ LN Equity
16	BLND LN Equity	49	KGF LN Equity	82	SMIN LN Equity
17	BLT LN Equity	50	LAND LN Equity	83	SN/ LN Equity
18	BNZL LN Equity	51	LGEN LN Equity	84	SRP LN Equity
19	BP/ LN Equity	52	LLOY LN Equity	85	SSE LN Equity
20	BRBY LN Equity	53	LSE LN Equity	86	STAN LN Equity
21	BSY LN Equity	54	MGGT LN Equity	87	SVT LN Equity
22	BT/A LN Equity	55	MKS LN Equity	88	TATE LN Equity
23	CCL LN Equity	56	MRO LN Equity	89	TLW LN Equity
24	CNA LN Equity	57	MRW LN Equity	90	TSCO LN Equity
25	CPG LN Equity	58	NG/ LN Equity	91	TT/ LN Equity
26	CPI LN Equity	59	NXT LN Equity	92	ULVR LN Equity
27	CRDA LN Equity	60	OML LN Equity	93	UU/ LN Equity
28	CRH LN Equity	61	PFC LN Equity	94	VED LN Equity
29	DGE LN Equity	62	POLY LN Equity	95	VOD LN Equity
30	ENRC LN Equity	63	PRU LN Equity	96	WEIR LN Equity
31	EVR LN Equity	64	PSON LN Equity	97	WG/ LN Equity
32	EXPN LN Equity	65	RB/ LN Equity	98	WMH LN Equity
33	EZJ LN Equity	66	RBS LN Equity	99	WOS LN Equity
				100	WPP LN Equity
				101	WTB LN Equity

Appendix O

European Stocks

Here we present all 250 stocks used in our MN Stocks model:

Table O.1: Largest European stocks

Number	Ticker	Number	Ticker
1	A2A IM Equity	26	MT NA Equity
2	ABG SM Equity	27	ASML NA Equity
3	ABE SM Equity	28	ATL IM Equity
4	ANA SM Equity	29	ATO FP Equity
5	AC FP Equity	30	AGL IM Equity
6	ACE IM Equity	31	CS FP Equity
7	ACX SM Equity	32	CRG IM Equity
8	ACKB BB Equity	33	BMPS IM Equity
9	ACS SM Equity	34	PMI IM Equity
10	ADS GR Equity	35	BPI PL Equity
11	ADP FP Equity	36	BCP PL Equity
12	AGN NA Equity	37	BES PL Equity
13	AGS SM Equity	38	BP IM Equity
14	AH NA Equity	39	POP SM Equity
15	AF FP Equity	40	SAB SM Equity
16	AI FP Equity	41	SAN SM Equity
17	AKZA NA Equity	42	BVA SM Equity
18	ALB SM Equity	43	BTO SM Equity
19	ALU FP Equity	44	BKIR ID Equity
20	ALV GR Equity	45	BKT SM Equity
21	ALBK ID Equity	46	BAS GR Equity
22	ALPHA GA Equity	47	BAYN GR Equity
23	ALO FP Equity	48	BMW GR Equity
24	ANDR AV Equity	49	BBVA SM Equity
25	ABI BB Equity	50	BEI GR Equity

Table O.2: Largest European stocks

Number	Ticker	Number	Ticker
51	BEKB BB Equity	76	CBK GR Equity
52	BELG BB Equity	77	CON GR Equity
53	BB FP Equity	78	ACA FP Equity
54	BIM FP Equity	79	CVAL IM Equity
55	BNP FP Equity	80	CRH ID Equity
56	BME SM Equity	81	CRI SM Equity
57	GBB FP Equity	82	DAI GR Equity
58	EN FP Equity	83	BN FP Equity
59	BRI PL Equity	84	DSY FP Equity
60	BVI FP Equity	85	DELB BB Equity
61	BZU IM Equity	86	DBK GR Equity
62	CPR IM Equity	87	DB1 GR Equity
63	CAP FP Equity	88	LHA GR Equity
64	CA FP Equity	89	DPW GR Equity
65	CO FP Equity	90	DPB GR Equity
66	GCO SM Equity	91	DTE GR Equity
67	CLS1 GR Equity	92	DEXB BB Equity
68	CEP SM Equity	93	DSM NA Equity
69	CDI FP Equity	94	EOAN GR Equity
70	CMA FP Equity	95	EAD FP Equity
71	CPR PL Equity	96	EVA SM Equity
72	NAT BB Equity	97	EDF FP Equity
73	CNP FP Equity	98	EEN FP Equity
74	EEEK GA Equity	99	EDN IM Equity
75	COLR BB Equity	100	EDP PL Equity

Table O.3: Largest European stocks

Number	Ticker	Number	Ticker
101	EDPR PL Equity	126	FTE FP Equity
102	EUROB GA Equity	127	FRA GR Equity
103	FGR FP Equity	128	FME GR Equity
104	ELN ID Equity	129	FRE GR Equity
105	ELI1V FH Equity	130	GALP PL Equity
106	ENG SM Equity	131	GAM SM Equity
107	ELE SM Equity	132	GAS SM Equity
108	ENEL IM Equity	133	GSZ FP Equity
109	ENI IM Equity	134	G1A GR Equity
110	ERA FP Equity	135	GTO FP Equity
111	ERG IM Equity	136	G IM Equity
112	EBS AV Equity	137	GA FP Equity
113	EI FP Equity	138	GRF SM Equity
114	ELE FP Equity	139	GBLB BB Equity
115	RF FP Equity	140	HHFA GR Equity
116	ETL FP Equity	141	HNR1 GR Equity
117	EVN AV Equity	142	HEI GR Equity
118	EXO IM Equity	143	HEIO NA Equity
119	FCC SM Equity	144	HEIA NA Equity
120	FER SM Equity	145	ELPE GA Equity
121	F IM Equity	146	HTO GA Equity
122	FIE GR Equity	147	HER IM Equity
123	FNC IM Equity	148	RMS FP Equity
124	FORB BB Equity	149	HOT GR Equity
125	FUM1V FH Equity	150	IBR SM Equity

Table O.4: Largest European stocks

Number	Ticker	Number	Ticker
151	IBE SM Equity	176	LUX IM Equity
152	IBLA SM Equity	177	MC FP Equity
153	ILD FP Equity	178	MMT FP Equity
154	NK FP Equity	179	MAN GR Equity
155	IEA AV Equity	180	MAP SM Equity
156	ITX SM Equity	181	MIG GA Equity
157	IDR SM Equity	182	MS IM Equity
158	IFX GR Equity	183	MB IM Equity
159	ISP IM Equity	184	MED IM Equity
160	IPN FP Equity	185	MRK GR Equity
161	IT IM Equity	186	MEO GR Equity
162	DEC FP Equity	187	MEO1V FH Equity
163	JMT PL Equity	188	ML FP Equity
164	SDF GR Equity	189	MOBB BB Equity
165	KBC BB Equity	190	MUV2 GR Equity
166	KYG ID Equity	191	KN FP Equity
167	KNEBV FH Equity	192	ETE GA Equity
168	KPN NA Equity	193	NEO FP Equity
169	OR FP Equity	194	NES1V FH Equity
170	LG FP Equity	195	NOK1V FH Equity
171	MMB FP Equity	196	NRE1V FH Equity
172	LXS GR Equity	197	VER AV Equity
173	LR FP Equity	198	POST AV Equity
174	LIN GR Equity	199	OMV AV Equity
175	LTO IM Equity	200	OPAP GA Equity

Table O.5: Largest European stocks

Number	Ticker	Number	Ticker
201	OUT1V FH Equity	226	RWE GR Equity
202	PAJ FP Equity	227	RYA ID Equity
203	PLT IM Equity	228	SYV SM Equity
204	RI FP Equity	229	SAF FP Equity
205	UG FP Equity	230	SGO FP Equity
206	PHIA NA Equity	231	SPM IM Equity
207	TPEIR GA Equity	232	SZG GR Equity
208	POH1S FH Equity	233	SAMAS FH Equity
209	PTC PL Equity	234	SAN FP Equity
210	PP FP Equity	235	SAA1V FH Equity
211	PSG SM Equity	236	SAP GR Equity
212	PRY IM Equity	237	SRS IM Equity
213	PPC GA Equity	238	SBMO NA Equity
214	PUB FP Equity	239	SU FP Equity
215	PUM GR Equity	240	SCR FP Equity
216	RIBH AV Equity	241	SK FP Equity
217	RAND NA Equity	242	SIE GR Equity
218	RTRKS FH Equity	243	SKYD GR Equity
219	REE SM Equity	244	S92 GR Equity
220	REN NA Equity	245	SRG IM Equity
221	REIN LX Equity	246	GLE FP Equity
222	RNO FP Equity	247	ARR FP Equity
223	REP SM Equity	248	SW FP Equity
224	RXL FP Equity	249	SOF BB Equity
225	RHK GR Equity	250	SOW GR Equity

Appendix P

Sector Categorisation of Euro denominated Stocks for MN Pairs

Table P.1: EU stocks sector classification

Ticker	Name	Sector	Sub Sector
EAD FP Equity	European Aeronautic Defence and Space Co NV	Aerospace & Defense	Aerospace & Defense
FNC IM Equity	Finmeccanica SpA	Aerospace & Defense	Aerospace & Defense
SAF FP Equity	Safran SA	Aerospace & Defense	Aerospace & Defense
HO FP Equity	Thales SA	Aerospace & Defense	Aerospace & Defense

TABLE P.2: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
BMW GR Equity	Bayerische Motoren Werke AG	Automobiles & Parts	Automobile Manufacturers
DAI GR Equity	Daimler AG	Automobiles & Parts	Automobile Manufacturers
F IM Equity	Fiat SpA	Automobiles & Parts	Automobile Manufacturers
UG FP Equity	Peugeot SA	Automobiles & Parts	Automobile Manufacturers
RNO FP Equity	Renault SA	Automobiles & Parts	Automobile Manufacturers
VOW GR Equity	Volkswagen AG	Automobiles & Parts	Automobile Manufacturers
CON GR Equity	Continental AG	Automobiles & Parts	Tires & Rubber
ML FP Equity	Compagnie Generale des Etablissements Michelin	Automobiles & Parts	Tires & Rubber
NRE1V FH Equity	Nokian Renkaat OYJ	Automobiles & Parts	Tires & Rubber

TABLE P.3: EU stocks sector classification

Ticker	Name	Sector	Sub Sector
ALBK ID Equity	Allied Irish Banks PLC	Banks	Diversified Banks
CRG IM Equity	Banca Carige SpA	Banks	Diversified Banks
BMPS IM Equity	Banca Monte dei Paschi di Siena SpA	Banks	Diversified Banks
PMI IM Equity	Banca Popolare di Milano Scarl	Banks	Diversified Banks
BPI PL Equity	Banco BPI SA	Banks	Diversified Banks
BCP PL Equity	Banco Comercial Portugues SA	Banks	Diversified Banks
BES PL Equity	Banco Espirito Santo SA	Banks	Diversified Banks
BP IM Equity	Banco Popolare SC	Banks	Diversified Banks
POP SM Equity	Banco Popular Espanol SA	Banks	Diversified Banks
SAB SM Equity	Banco de Sabadell SA	Banks	Diversified Banks
SAN SM Equity	Banco Santander SA	Banks	Diversified Banks
BKIR ID Equity	Governor & Co of the Bank of Ireland/The	Banks	Diversified Banks
BKT SM Equity	Bankinter SA	Banks	Diversified Banks
BBVA SM Equity	Banco Bilbao Vizcaya Argentaria SA	Banks	Diversified Banks
BNP FP Equity	BNP Paribas	Banks	Diversified Banks
CBK GR Equity	Commerzbank AG	Banks	Diversified Banks
ACA FP Equity	Credit Agricole SA	Banks	Diversified Banks
DBK GR Equity	Deutsche Bank AG	Banks	Diversified Capital Markets
DEXB BB Equity	Dexia SA	Banks	Diversified Banks
EUROB GA Equity	EFG Eurobank Ergasias SA	Banks	Diversified Banks
EBS AV Equity	Erste Group Bank AG	Banks	Diversified Banks
ISP IM Equity	Intesa Sanpaolo SpA	Banks	Diversified Banks
KBC BB Equity	KBC Groep NV	Banks	Diversified Banks
MB IM Equity	Mediobanca SpA	Banks	Investment Banking & Brokerage
KN FP Equity	Natixis	Banks	Diversified Banks
GLE FP Equity	Societe Generale	Banks	Diversified Banks
UBI IM Equity	Unione di Banche Italiane SCPA	Banks	Diversified Banks
UCG IM Equity	UniCredit SpA	Banks	Diversified Banks
BTO SM Equity	Banco Espanol de Credito SA	Banks	Diversified Banks
DPB GR Equity	Deutsche Postbank AG	Banks	Diversified Banks

Table P.4: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
ABI BB Equity	Anheuser-Busch InBev NV	Beverages	Brewers
HEIO NA Equity	Heineken Holding NV	Beverages	Brewers
HEIA NA Equity	Heineken NV	Beverages	Brewers
RI FP Equity	Pernod-Ricard SA	Beverages	Distillers & Vintners
CPR IM Equity	Davide Campari-Milano SpA	Beverages	Distillers & Vintners

Table P.5: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
AKZA NA Equity	Akzo Nobel NV	Chemicals	Diversified Chemicals
BAS GR Equity	BASF SE	Chemicals	Diversified Chemicals
SOLB BB Equity	Solvay SA	Chemicals	Diversified Chemicals
LXS GR Equity	Lanxess AG	Chemicals	Diversified Chemicals
UMI BB Equity	Umicore	Chemicals	Specialty Chemicals
WCH GR Equity	Wacker Chemie AG	Chemicals	Specialty Chemicals
DSM NA Equity	Koninklijke DSM NV	Chemicals	Specialty Chemicals

Table P.6: EU stocks sector classification

Ticker	Name	Sector	Sub Sector
LIN GR Equity	Linde AG	Chemicals	Industrial Gases
AI FP Equity	Air Liquide SA	Chemicals	Industrial Gases
BAYN GR Equity	Bayer AG	Chemicals	Pharmaceuticals
SDF GR Equity	K+S AG	Chemicals	Fertilizers & Agricultural Che

Table P.7: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
ABG SM Equity	Abengoa SA	Construction & Materials	Construction & Engineering
ACS SM Equity	ACS Actividades de Construccion y Servicios SA	Construction & Materials	Construction & Engineering
EN FP Equity	Bouygues SA	Construction & Materials	Construction & Engineering
FGR FP Equity	Eiffage SA	Construction & Materials	Construction & Engineering
FCC SM Equity	Fomento de Construcciones y Contratas SA	Construction & Materials	Construction & Engineering
HOT GR Equity	Hochtief AG	Construction & Materials	Construction & Engineering
SYV SM Equity	Sacyr Vallehermoso SA	Construction & Materials	Construction & Engineering
STR AV Equity	Strabag SE	Construction & Materials	Construction & Engineering
DG FP Equity	Vinci SA	Construction & Materials	Construction & Engineering

Table P.8: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
BZU IM Equity	Buzzi Unicem SpA	Construction & Materials	Construction Materials
CPR PL Equity	Cimpor Cimentos de Portugal SGPS SA	Construction & Materials	Construction Materials
CRH ID Equity	CRH PLC	Construction & Materials	Construction Materials
HEI GR Equity	HeidelbergCement AG	Construction & Materials	Construction Materials
IT IM Equity	Italcementi SpA	Construction & Materials	Construction Materials
LG FP Equity	Lafarge SA	Construction & Materials	Construction Materials

TABLE P.9: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
LR FP Equity	Legrand SA	Electronic & Electrical Equipm	Electrical Components & Equipm
PRY IM Equity	Prysmian SpA	Electronic & Electrical Equipm	Electrical Components & Equipm
SU FP Equity	Schneider Electric SA	Electronic & Electrical Equipm	Electrical Components & Equipm

TABLE P.10: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
ACKB BB Equity	Ackermans & van Haaren NV	Financial Services	Multi-Sector Holdings
ALB SM Equity	Corp Financiera Alba	Financial Services	Multi-Sector Holdings
NAT BB Equity	Cie Nationale a Portefeuille	Financial Services	Multi-Sector Holdings
CRI SM Equity	Criteria Caixacorp SA	Financial Services	Multi-Sector Holdings
RF FP Equity	Eurazeo	Financial Services	Multi-Sector Holdings
EXO IM Equity	Exor SpA	Financial Services	Multi-Sector Holdings
GBLB BB Equity	Groupe Bruxelles Lambert SA	Financial Services	Multi-Sector Holdings
SOF BB Equity	Sofina SA	Financial Services	Multi-Sector Holdings

TABLE P.11: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
DTE GR Equity	Deutsche Telekom AG	Mobile Telecommunications	Integrated Telecommunication S
BELG BB Equity	Belgacom SA	Fixed Line Telecommunications	Integrated Telecommunication S
ELI1V FH Equity	Elisa OYJ	Fixed Line Telecommunications	Integrated Telecommunication S
FTE FP Equity	France Telecom SA	Fixed Line Telecommunications	Integrated Telecommunication S
KPN NA Equity	Koninklijke KPN NV	Fixed Line Telecommunications	Integrated Telecommunication S
PTC PL Equity	Portugal Telecom SGPS SA	Fixed Line Telecommunications	Integrated Telecommunication S
TIT IM Equity	Telecom Italia SpA	Fixed Line Telecommunications	Integrated Telecommunication S
TEF SM Equity	Telefonica SA	Fixed Line Telecommunications	Integrated Telecommunication S
TKA AV Equity	Telekom Austria AG	Fixed Line Telecommunications	Integrated Telecommunication S

TABLE P.12: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
AH NA Equity	Koninklijke Ahold NV	Food & Drug Retailers	Food Retail
COLR BB Equity	Colruyt SA	Food & Drug Retailers	Food Retail
DELB BB Equity	Delhaize Group SA	Food & Drug Retailers	Food Retail
JMT PL Equity	Jeronimo Martins SGPS SA	Food & Drug Retailers	Food Retail
CO FP Equity	Casino Guichard Perrachon SA	Food & Drug Retailers	Food Retail
CA FP Equity	Carrefour SA	Food & Drug Retailers	Hypermarkets & Super Centers
CLS1 GR Equity	Celesio AG	Food & Drug Retailers	Health Care Distributors

TABLE P.13: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
BN FP Equity	Danone	Food Producers	Packaged Foods & Meats
EVA SM Equity	Ebro Puleva SA	Food Producers	Packaged Foods & Meats
KYG ID Equity	Kerry Group PLC	Food Producers	Packaged Foods & Meats
PLT IM Equity	Parmalat SpA	Food Producers	Packaged Foods & Meats

Table P.14: EU stocks sector classification

Ticker	Name	Sector	Sub Sector
IBE SM Equity	Iberdrola SA	Electricity	Electric Utilities
VER AV Equity	Verbund - Oesterreichische Elektrizitaetswirtschafts AG	Electricity	Electric Utilities
REE SM Equity	Red Electrica Corp SA	Electricity	Electric Utilities
TRN IM Equity	Terna Rete Elettrica Nazionale SpA	Electricity	Electric Utilities
ELE SM Equity	Endesa SA	Electricity	Electric Utilities
ENEL IM Equity	Enel SpA	Electricity	Electric Utilities
FUM1V FH Equity	Fortum Oyj	Electricity	Electric Utilities
EDF FP Equity	EDF SA	Electricity	Electric Utilities
EDP PL Equity	EDP - Energias de Portugal SA	Electricity	Electric Utilities
EOAN GR Equity	E.ON AG	Gas, Water & Multiutilities	Electric Utilities
EVN AV Equity	EVN AG	Gas, Water & Multiutilities	Electric Utilities
ENG SM Equity	Enagas	Gas, Water & Multiutilities	Gas Utilities
SRG IM Equity	Snam Rete Gas SpA	Gas, Water & Multiutilities	Gas Utilities
GAS SM Equity	Gas Natural SDG SA	Gas, Water & Multiutilities	Gas Utilities
A2A IM Equity	A2A SpA	Electricity	Multi-Utilities
ACE IM Equity	ACEA SpA	Electricity	Multi-Utilities
GSZ FP Equity	GDF Suez	Gas, Water & Multiutilities	Multi-Utilities
HER IM Equity	Hera SpA	Gas, Water & Multiutilities	Multi-Utilities
RWE GR Equity	RWE AG	Gas, Water & Multiutilities	Multi-Utilities
VIE FP Equity	Veolia Environnement	Gas, Water & Multiutilities	Multi-Utilities
EDN IM Equity	Edison SpA	Gas, Water & Multiutilities	Independent Power Producers &
EEN FP Equity	EDF Energies Nouvelles SA	Electricity	Independent Power Producers &
EDPR PL Equity	EDP Renovaveis SA	Electricity	Independent Power Producers &
IBR SM Equity	Iberdrola Renovables SA	Electricity	Independent Power Producers &
AGS SM Equity	Sociedad General de Aguas de Barcelona SA	Gas, Water & Multiutilities	Water Utilities

TABLE P.15: EU stocks sector classification

Ticker	Name	Sector	Sub Sector
G1A GR Equity	GEA Group AG	General Industrials	Industrial Machinery
MEO1V FH Equity	Metso Oyj	Industrial Engineering	Industrial Machinery
VK FP Equity	Vallourec SA	Industrial Engineering	Industrial Machinery
WRT1V FH Equity	Wartsila Oyj	Industrial Engineering	Industrial Machinery
ZOT SM Equity	Zardoya Otis SA	Industrial Engineering	Industrial Machinery
ANDR AV Equity	Andritz AG	Industrial Engineering	Industrial Machinery
KNEBV FH Equity	Kone OYJ	Industrial Engineering	Industrial Machinery
SIE GR Equity	Siemens AG	General Industrials	Industrial Conglomerates
MAN GR Equity	MAN SE	Industrial Engineering	Construction & Farm Machinery
ALO FP Equity	Alstom SA	Industrial Engineering	Heavy Electrical Equipment

Table P.16: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
BIM FP Equity	BioMerieux	Health Care Equipment & Servic	Health Care Equipment
EI FP Equity	Cie Generale d'Optique Essilor International SA	Health Care Equipment & Servic	Health Care Supplies
FME GR Equity	Fresenius Medical Care AG & Co KGaA	Health Care Equipment & Servic	Health Care Services
FRE GR Equity	Fresenius SE	Health Care Equipment & Servic	Health Care Equipment
RHK GR Equity	Rhoen Klinikum AG	Health Care Equipment & Servic	Health Care Facilities

Table P.17: EU stocks sector classification

Ticker	Name	Sector	Sub Sector
TKA GR Equity	ThyssenKrupp AG	General Industrials	Steel
ACX SM Equity	Acerinox SA	Industrial Metals & Mining	Steel
MT NA Equity	ArcelorMittal	Industrial Metals & Mining	Steel
OUT1V FH Equity	Outokumpu OYJ	Industrial Metals & Mining	Steel
RTRKS FH Equity	Rautaruukki OYJ	Industrial Metals & Mining	Steel
SZG GR Equity	Salzgitter AG	Industrial Metals & Mining	Steel
VOE AV Equity	Voestalpine AG	Industrial Metals & Mining	Steel
TEN IM Equity	Tenaris SA	Industrial Metals & Mining	Oil & Gas Equipment & Services
ERA FP Equity	Eramet	Industrial Metals & Mining	Diversified Metals & Mining
ARR FP Equity	Societe Des Autoroutes Paris-Rhin-Rhone	Industrial Transportation	Highways & Railtracks
ABE SM Equity	Abertis Infraestructuras SA	Industrial Transportation	Highways & Railtracks
ATL IM Equity	Atlantia SpA	Industrial Transportation	Highways & Railtracks
BRI PL Equity	Brisa Auto-Estradas de Portugal SA	Industrial Transportation	Highways & Railtracks
FER SM Equity	Ferrovial SA	Industrial Transportation	Highways & Railtracks
DPW GR Equity	Deutsche Post AG	Industrial Transportation	Air Freight & Logistics
TNT NA Equity	TNT NV	Industrial Transportation	Air Freight & Logistics
POST AV Equity	Oesterreichische Post AG	Industrial Transportation	Air Freight & Logistics
HHFA GR Equity	Hamburger Hafen und Logistik AG	Industrial Transportation	Marine Ports & Services
VPK NA Equity	Koninklijke Vopak NV	Industrial Transportation	Marine Ports & Services

TABLE P.18: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
PHIA NA Equity	Koninklijke Philips Electronics NV	Leisure Goods	Industrial Conglomerates
UBI FP Equity	UBISOFT Entertainment	Leisure Goods	Home Entertainment Software

Table P.19: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
AGN NA Equity	Aegon NV	Life Insurance	Life & Health Insurance
CNP FP Equity	CNP Assurances	Life Insurance	Life & Health Insurance
MED IM Equity	Mediolanum SpA	Life Insurance	Life & Health Insurance
FORB BB Equity	Fortis	Life Insurance	Multi-line Insurance
ALV GR Equity	Allianz SE	Nonlife Insurance	Multi-line Insurance
CS FP Equity	AXA SA	Nonlife Insurance	Multi-line Insurance
GCO SM Equity	Grupo Catalana Occidente SA	Nonlife Insurance	Multi-line Insurance
G IM Equity	Assicurazioni Generali SpA	Nonlife Insurance	Multi-line Insurance
MAP SM Equity	Mapfre SA	Nonlife Insurance	Multi-line Insurance
SAMAS FH Equity	Sampo Oyj	Nonlife Insurance	Multi-line Insurance
MUV2 GR Equity	Muenchener Rueckversicherungs AG	Nonlife Insurance	Reinsurance
SCR FP Equity	SCOR SE	Nonlife Insurance	Reinsurance
HNR1 GR Equity	Hannover Rueckversicherung AG	Nonlife Insurance	Reinsurance
VIG AV Equity	Vienna Insurance Group	Nonlife Insurance	Multi-line Insurance
ELE FP Equity	Euler Hermes SA	Nonlife Insurance	Property & Casualty Insurance

TABLE P.20: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
CEP SM Equity	Cia Espanola de Petroleos SA	Oil & Gas Producers	Integrated Oil & Gas
ENI IM Equity	ENI SpA	Oil & Gas Producers	Integrated Oil & Gas
GALP PL Equity	Galp Energia SGPS SA	Oil & Gas Producers	Integrated Oil & Gas
OMV AV Equity	OMV AG	Oil & Gas Producers	Integrated Oil & Gas
REP SM Equity	Repsol YPF SA	Oil & Gas Producers	Integrated Oil & Gas
FP FP Equity	Total SA	Oil & Gas Producers	Integrated Oil & Gas
ERG IM Equity	ERG SpA	Oil & Gas Producers	Oil & Gas Refining & Marketing
NES1V FH Equity	Neste Oil OYJ	Oil & Gas Producers	Oil & Gas Refining & Marketing
SRS IM Equity	Saras SpA	Oil & Gas Producers	Oil & Gas Refining & Marketing

TABLE P.21: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
ETL FP Equity	Eutelsat Communications	Media	Cable & Satellite
SKYD GR Equity	Sky Deutschland AG	Media	Cable & Satellite
DEC FP Equity	JC Decaux SA	Media	Advertising
PUB FP Equity	Publicis Groupe SA	Media	Advertising
MMB FP Equity	Lagardere SCA	Media	Publishing
PAJ FP Equity	PagesJaunes Groupe	Media	Publishing
REN NA Equity	Reed Elsevier NV	Media	Publishing
SAA1V FH Equity	Sanoma Oyj	Media	Publishing
WKL NA Equity	Wolters Kluwer NV	Media	Publishing
MMT FP Equity	M6-Metropole Television	Media	Broadcasting
MS IM Equity	Mediaset SpA	Media	Broadcasting
TL5 SM Equity	Gestevision Telecinco SA	Media	Broadcasting
TFI FP Equity	Societe Television Francaise 1	Media	Broadcasting
TNET BB Equity	Telenet Group Holding NV	Media	Alternative Carriers
VIV FP Equity	Vivendi SA	Media	Movies & Entertainment

Table P.22: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
GBB FP Equity	Bourbon SA	Oil Equipment, Services & Dist	Oil & Gas Equipment & Services
GA FP Equity	Cie Generale de Geophysique-Veritas	Oil Equipment, Services & Dist	Oil & Gas Equipment & Services
SPM IM Equity	Saipem SpA	Oil Equipment, Services & Dist	Oil & Gas Equipment & Services
SBMO NA Equity	SBM Offshore NV	Oil Equipment, Services & Dist	Oil & Gas Equipment & Services
TEC FP Equity	Technip SA	Oil Equipment, Services & Dist	Oil & Gas Equipment & Services

TABLE P.23: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
ADS GR Equity	Adidas AG	Personal Goods	Apparel, Accessories & Luxury
CDI FP Equity	Christian Dior SA	Personal Goods	Apparel, Accessories & Luxury
RMS FP Equity	Hermes International	Personal Goods	Apparel, Accessories & Luxury
LUX IM Equity	Luxottica Group SpA	Personal Goods	Apparel, Accessories & Luxury
MC FP Equity	LVMH Moet Hennessy Louis Vuitton SA	Personal Goods	Apparel, Accessories & Luxury
BEI GR Equity	Beiersdorf AG	Personal Goods	Personal Products
OR FP Equity	L'Oreal SA	Personal Goods	Personal Products
PUM GR Equity	Puma AG Rudolf Dassler Sport	Personal Goods	Footwear

Table P.24: EU stocks sector classification

Ticker	Name	Sector	Sub Sector
ELN ID Equity	Elan Corp PLC	Pharmaceuticals & Biotechnolog	Pharmaceuticals
IPN FP Equity	Ipsen SA	Pharmaceuticals & Biotechnolog	Pharmaceuticals
MRK GR Equity	Merck KGaA	Pharmaceuticals & Biotechnolog	Pharmaceuticals
SAN FP Equity	Sanofi-Aventis SA	Pharmaceuticals & Biotechnolog	Pharmaceuticals
UCB BB Equity	UCB SA	Pharmaceuticals & Biotechnolog	Pharmaceuticals
GRF SM Equity	Grifols SA	Pharmaceuticals & Biotechnolog	Biotechnology

TABLE P.25: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
ATO FP Equity	Atos Origin SA	Software & Computer Services	IT Consulting & Other Services
CAP FP Equity	Cap Gemini SA	Software & Computer Services	IT Consulting & Other Services
IDR SM Equity	Indra Sistemas SA	Software & Computer Services	IT Consulting & Other Services
DSY FP Equity	Dassault Systemes SA	Software & Computer Services	Application Software
SAP GR Equity	SAP AG	Software & Computer Services	Application Software
ILD FP Equity	Iliad SA	Software & Computer Services	Alternative Carriers
SOW GR Equity	Software AG	Software & Computer Services	Systems Software
UTDI GR Equity	United Internet AG	Software & Computer Services	Internet Software & Services

TABLE P.26: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
BVI FP Equity	Bureau Veritas SA	Support Services	Research and Consulting Servic
PSG SM Equity	Prosegur Cia de Seguridad SA	Support Services	Security & Alarm Services
RAND NA Equity	Randstad Holding NV	Support Services	Human Resource & Employment Se
SEV FP Equity	Suez Environnement Co	Support Services	Multi-Utilities

Table P.27: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
ALU FP Equity	Alcatel-Lucent	Technology Hardware & Equipmen	Communications Equipment
NOK1V FH Equity	Nokia OYJ	Technology Hardware & Equipmen	Communications Equipment
IFX GR Equity	Infineon Technologies AG	Technology Hardware & Equipmen	Semiconductors
STM FP Equity	STMicroelectronics NV	Technology Hardware & Equipmen	Semiconductors
ASML NA Equity	ASML Holding NV	Technology Hardware & Equipmen	Semiconductor Equipment
NEO FP Equity	Neopost SA	Technology Hardware & Equipmen	Office Electronics

Table P.28: EU stocks sector classification

Ticker	**Name**	**Sector**	**Sub Sector**
AF FP Equity	Air France-KLM	Travel & Leisure	Airlines
LHA GR Equity	Deutsche Lufthansa AG	Travel & Leisure	Airlines
IBLA SM Equity	Iberia Lineas Aereas de Espana SA	Travel & Leisure	Airlines
RYA ID Equity	Ryanair Holdings PLC	Travel & Leisure	Airlines
AGL IM Equity	Autogrill SpA	Travel & Leisure	Restaurants
SW FP Equity	Sodexo	Travel & Leisure	Restaurants
LTO IM Equity	Lottomatica SpA	Travel & Leisure	Casinos & Gaming
OPAP GA Equity	OPAP SA	Travel & Leisure	Casinos & Gaming
AC FP Equity	Accor SA	Travel & Leisure	Hotels, Resorts & Cruise Lines

Appendix Q

List of EU all Stock Pairs

This list all the sector pairs before they get short listed.

Table Q.1: EU stock pairs

EAD FP Equity	FNC IM Equity
EAD FP Equity	SAF FP Equity
EAD FP Equity	HO FP Equity
FNC IM Equity	SAF FP Equity
FNC IM Equity	HO FP Equity
SAF FP Equity	HO FP Equity
BMW GR Equity	DAI GR Equity
BMW GR Equity	F IM Equity
BMW GR Equity	UG FP Equity
BMW GR Equity	RNO FP Equity
BMW GR Equity	VOW GR Equity
DAI GR Equity	F IM Equity
DAI GR Equity	UG FP Equity
DAI GR Equity	RNO FP Equity
DAI GR Equity	VOW GR Equity
F IM Equity	UG FP Equity
F IM Equity	RNO FP Equity
F IM Equity	VOW GR Equity
UG FP Equity	RNO FP Equity
UG FP Equity	VOW GR Equity
RNO FP Equity	VOW GR Equity
CON GR Equity	ML FP Equity
GAM SM Equity	S92 GR Equity
GAM SM Equity	SWV GR Equity
S92 GR Equity	SWV GR Equity
HEIO NA Equity	HEIA NA Equity
HEIO NA Equity	RI FP Equity
HEIO NA Equity	CPR IM Equity
HEIA NA Equity	RI FP Equity
HEIA NA Equity	CPR IM Equity

Table Q.2: EU stock pairs

RI FP Equity	CPR IM Equity
AKZA NA Equity	BAS GR Equity
AKZA NA Equity	LXS GR Equity
AKZA NA Equity	WCH GR Equity
AKZA NA Equity	DSM NA Equity
BAS GR Equity	LXS GR Equity
BAS GR Equity	WCH GR Equity
BAS GR Equity	DSM NA Equity
LXS GR Equity	WCH GR Equity
LXS GR Equity	DSM NA Equity
WCH GR Equity	DSM NA Equity
LIN GR Equity	AI FP Equity
LIN GR Equity	BAYN GR Equity
LIN GR Equity	SDF GR Equity
AI FP Equity	BAYN GR Equity
AI FP Equity	SDF GR Equity
BAYN GR Equity	SDF GR Equity
ABG SM Equity	ACS SM Equity
ABG SM Equity	EN FP Equity
ABG SM Equity	FGR FP Equity
ABG SM Equity	FCC SM Equity
ABG SM Equity	HOT GR Equity
ABG SM Equity	SYV SM Equity
ABG SM Equity	STR AV Equity
ABG SM Equity	DG FP Equity
ACS SM Equity	EN FP Equity
ACS SM Equity	FGR FP Equity
ACS SM Equity	FCC SM Equity
ACS SM Equity	HOT GR Equity
ACS SM Equity	SYV SM Equity

Table Q.3: EU stock pairs

ACS SM Equity	STR AV Equity
ACS SM Equity	DG FP Equity
EN FP Equity	FGR FP Equity
EN FP Equity	FCC SM Equity
EN FP Equity	HOT GR Equity
EN FP Equity	SYV SM Equity
EN FP Equity	STR AV Equity
EN FP Equity	DG FP Equity
FGR FP Equity	FCC SM Equity
FGR FP Equity	HOT GR Equity
FGR FP Equity	SYV SM Equity
FGR FP Equity	STR AV Equity
FGR FP Equity	DG FP Equity
FCC SM Equity	HOT GR Equity
FCC SM Equity	SYV SM Equity
FCC SM Equity	STR AV Equity
FCC SM Equity	DG FP Equity
HOT GR Equity	SYV SM Equity
HOT GR Equity	STR AV Equity
HOT GR Equity	DG FP Equity
SYV SM Equity	STR AV Equity
SYV SM Equity	DG FP Equity
STR AV Equity	DG FP Equity
BZU IM Equity	HEI GR Equity
BZU IM Equity	IT IM Equity
BZU IM Equity	LG FP Equity
HEI GR Equity	IT IM Equity
HEI GR Equity	LG FP Equity
IT IM Equity	LG FP Equity
SGO FP Equity	TRE SM Equity

Table Q.4: EU stock pairs

SGO FP Equity	ANA SM Equity
TRE SM Equity	ANA SM Equity
LR FP Equity	PRY IM Equity
LR FP Equity	SU FP Equity
PRY IM Equity	SU FP Equity
ALB SM Equity	CRI SM Equity
ALB SM Equity	RF FP Equity
ALB SM Equity	EXO IM Equity
CRI SM Equity	RF FP Equity
CRI SM Equity	EXO IM Equity
BME SM Equity	DB1 GR Equity
FTE FP Equity	KPN NA Equity
FTE FP Equity	TIT IM Equity
FTE FP Equity	TEF SM Equity
FTE FP Equity	TKA AV Equity
KPN NA Equity	TIT IM Equity
KPN NA Equity	TEF SM Equity
KPN NA Equity	TKA AV Equity
TIT IM Equity	TEF SM Equity
TIT IM Equity	TKA AV Equity
TEF SM Equity	TKA AV Equity
CA FP Equity	CLS1 GR Equity
AH NA Equity	CO FP Equity
CRG IM Equity	BMPS IM Equity
CRG IM Equity	PMI IM Equity
CRG IM Equity	BP IM Equity
CRG IM Equity	POP SM Equity
CRG IM Equity	SAB SM Equity
CRG IM Equity	SAN SM Equity
CRG IM Equity	BKT SM Equity

Table Q.5: EU stock pairs

CRG IM Equity	BBVA SM Equity
CRG IM Equity	BNP FP Equity
CRG IM Equity	CBK GR Equity
CRG IM Equity	ACA FP Equity
CRG IM Equity	DBK GR Equity
CRG IM Equity	EBS AV Equity
CRG IM Equity	ISP IM Equity
CRG IM Equity	MB IM Equity
CRG IM Equity	KN FP Equity
CRG IM Equity	GLE FP Equity
CRG IM Equity	UBI IM Equity
CRG IM Equity	UCG IM Equity
CRG IM Equity	BTO SM Equity
CRG IM Equity	DPB GR Equity
BMPS IM Equity	PMI IM Equity
BMPS IM Equity	BP IM Equity
BMPS IM Equity	POP SM Equity
BMPS IM Equity	SAB SM Equity
BMPS IM Equity	SAN SM Equity
BMPS IM Equity	BKT SM Equity
BMPS IM Equity	BBVA SM Equity
BMPS IM Equity	BNP FP Equity
BMPS IM Equity	CBK GR Equity
BMPS IM Equity	ACA FP Equity
BMPS IM Equity	DBK GR Equity
BMPS IM Equity	EUROB GA Equity
BMPS IM Equity	EBS AV Equity
BMPS IM Equity	ISP IM Equity
BMPS IM Equity	MB IM Equity
BMPS IM Equity	KN FP Equity

Table Q.6: EU stock pairs

BMPS IM Equity	GLE FP Equity
BMPS IM Equity	UBI IM Equity
BMPS IM Equity	UCG IM Equity
BMPS IM Equity	BTO SM Equity
BMPS IM Equity	DPB GR Equity
PMI IM Equity	BP IM Equity
PMI IM Equity	POP SM Equity
PMI IM Equity	SAB SM Equity
PMI IM Equity	SAN SM Equity
PMI IM Equity	BKT SM Equity
PMI IM Equity	BBVA SM Equity
PMI IM Equity	BNP FP Equity
PMI IM Equity	CBK GR Equity
PMI IM Equity	ACA FP Equity
PMI IM Equity	DBK GR Equity
PMI IM Equity	EUROB GA Equity
PMI IM Equity	EBS AV Equity
PMI IM Equity	ISP IM Equity
PMI IM Equity	MB IM Equity
PMI IM Equity	KN FP Equity
PMI IM Equity	GLE FP Equity
PMI IM Equity	UBI IM Equity
PMI IM Equity	UCG IM Equity
PMI IM Equity	BTO SM Equity
PMI IM Equity	DPB GR Equity
BP IM Equity	POP SM Equity
BP IM Equity	SAB SM Equity
BP IM Equity	SAN SM Equity
BP IM Equity	BKT SM Equity
BP IM Equity	BBVA SM Equity

Table Q.7: EU stock pairs

BP IM Equity	BNP FP Equity
BP IM Equity	CBK GR Equity
BP IM Equity	ACA FP Equity
BP IM Equity	DBK GR Equity
BP IM Equity	EUROB GA Equity
BP IM Equity	EBS AV Equity
BP IM Equity	ISP IM Equity
BP IM Equity	MB IM Equity
BP IM Equity	KN FP Equity
BP IM Equity	GLE FP Equity
BP IM Equity	UBI IM Equity
BP IM Equity	UCG IM Equity
BP IM Equity	BTO SM Equity
BP IM Equity	DPB GR Equity
POP SM Equity	SAB SM Equity
POP SM Equity	SAN SM Equity
POP SM Equity	BKT SM Equity
POP SM Equity	BBVA SM Equity
POP SM Equity	BNP FP Equity
POP SM Equity	CBK GR Equity
BP IM Equity	POP SM Equity
BP IM Equity	SAB SM Equity
BP IM Equity	SAN SM Equity
BP IM Equity	BKT SM Equity
BP IM Equity	BBVA SM Equity
BP IM Equity	BNP FP Equity
BP IM Equity	CBK GR Equity
BP IM Equity	ACA FP Equity
BP IM Equity	DBK GR Equity
BP IM Equity	EUROB GA Equity

Table Q.8: EU stock pairs

BP IM Equity	EBS AV Equity
BP IM Equity	ISP IM Equity
BP IM Equity	MB IM Equity
BP IM Equity	KN FP Equity
BP IM Equity	GLE FP Equity
BP IM Equity	UBI IM Equity
BP IM Equity	UCG IM Equity
BP IM Equity	BTO SM Equity
BP IM Equity	DPB GR Equity
POP SM Equity	SAB SM Equity
POP SM Equity	SAN SM Equity
POP SM Equity	BKT SM Equity
POP SM Equity	BBVA SM Equity
POP SM Equity	BNP FP Equity
POP SM Equity	CBK GR Equity
BBVA SM Equity	UBI IM Equity
BBVA SM Equity	UCG IM Equity
BBVA SM Equity	BTO SM Equity
BBVA SM Equity	DPB GR Equity
BNP FP Equity	CBK GR Equity
BNP FP Equity	ACA FP Equity
BNP FP Equity	DBK GR Equity
BNP FP Equity	EUROB GA Equity
BNP FP Equity	EBS AV Equity
BNP FP Equity	ISP IM Equity
BNP FP Equity	MB IM Equity
BNP FP Equity	KN FP Equity
BNP FP Equity	GLE FP Equity
BNP FP Equity	UBI IM Equity
BNP FP Equity	UCG IM Equity

Table Q.9: EU stock pairs

BNP FP Equity	BTO SM Equity
BNP FP Equity	DPB GR Equity
CBK GR Equity	ACA FP Equity
CBK GR Equity	DBK GR Equity
CBK GR Equity	EUROB GA Equity
CBK GR Equity	EBS AV Equity
CBK GR Equity	ISP IM Equity
CBK GR Equity	MB IM Equity
CBK GR Equity	KN FP Equity
CBK GR Equity	GLE FP Equity
CBK GR Equity	UBI IM Equity
CBK GR Equity	UCG IM Equity
CBK GR Equity	BTO SM Equity
CBK GR Equity	DPB GR Equity
ACA FP Equity	DBK GR Equity
ACA FP Equity	EUROB GA Equity
ACA FP Equity	EBS AV Equity
ACA FP Equity	ISP IM Equity
ACA FP Equity	KBC BB Equity
ACA FP Equity	MB IM Equity
ACA FP Equity	KN FP Equity
ACA FP Equity	GLE FP Equity
ACA FP Equity	UBI IM Equity
ACA FP Equity	UCG IM Equity
ACA FP Equity	BTO SM Equity
ACA FP Equity	DPB GR Equity
DBK GR Equity	EUROB GA Equity
DBK GR Equity	EBS AV Equity
DBK GR Equity	ISP IM Equity
DBK GR Equity	MB IM Equity

Table Q.10: EU stock pairs

DBK GR Equity	KN FP Equity
DBK GR Equity	GLE FP Equity
DBK GR Equity	UBI IM Equity
DBK GR Equity	UCG IM Equity
DBK GR Equity	BTO SM Equity
DBK GR Equity	DPB GR Equity
EUROB GA Equity	EBS AV Equity
EUROB GA Equity	ISP IM Equity
EUROB GA Equity	MB IM Equity
EUROB GA Equity	KN FP Equity
EUROB GA Equity	GLE FP Equity
EUROB GA Equity	UBI IM Equity
EUROB GA Equity	UCG IM Equity
EUROB GA Equity	BTO SM Equity
EUROB GA Equity	DPB GR Equity
EBS AV Equity	ISP IM Equity
EBS AV Equity	MB IM Equity
EBS AV Equity	KN FP Equity
EBS AV Equity	GLE FP Equity
EBS AV Equity	UBI IM Equity
EBS AV Equity	UCG IM Equity
EBS AV Equity	BTO SM Equity
EBS AV Equity	DPB GR Equity
ISP IM Equity	MB IM Equity
ISP IM Equity	KN FP Equity
ISP IM Equity	GLE FP Equity
ISP IM Equity	UBI IM Equity
ISP IM Equity	UCG IM Equity
ISP IM Equity	BTO SM Equity
ISP IM Equity	DPB GR Equity

Table Q.11: EU stock pairs

MB IM Equity	KN FP Equity
MB IM Equity	GLE FP Equity
MB IM Equity	UBI IM Equity
MB IM Equity	UCG IM Equity
MB IM Equity	BTO SM Equity
MB IM Equity	DPB GR Equity
KN FP Equity	GLE FP Equity
KN FP Equity	UBI IM Equity
KN FP Equity	UCG IM Equity
KN FP Equity	BTO SM Equity
KN FP Equity	DPB GR Equity
GLE FP Equity	UBI IM Equity
GLE FP Equity	UCG IM Equity
GLE FP Equity	BTO SM Equity
GLE FP Equity	DPB GR Equity
UBI IM Equity	UCG IM Equity
UBI IM Equity	BTO SM Equity
UBI IM Equity	DPB GR Equity
UCG IM Equity	BTO SM Equity
UCG IM Equity	DPB GR Equity
BTO SM Equity	DPB GR Equity
BN FP Equity	EVA SM Equity
BN FP Equity	PLT IM Equity
EVA SM Equity	PLT IM Equity
IBE SM Equity	VER AV Equity
IBE SM Equity	REE SM Equity
IBE SM Equity	TRN IM Equity
IBE SM Equity	ELE SM Equity
IBE SM Equity	ENEL IM Equity
IBE SM Equity	EDF FP Equity

Table Q.12: EU stock pairs

IBE SM Equity	EOAN GR Equity
IBE SM Equity	EVN AV Equity
VER AV Equity	REE SM Equity
VER AV Equity	TRN IM Equity
VER AV Equity	ELE SM Equity
VER AV Equity	ENEL IM Equity
VER AV Equity	EDF FP Equity
VER AV Equity	EOAN GR Equity
VER AV Equity	EVN AV Equity
REE SM Equity	TRN IM Equity
REE SM Equity	ELE SM Equity
REE SM Equity	ENEL IM Equity
REE SM Equity	EDF FP Equity
REE SM Equity	EOAN GR Equity
REE SM Equity	EVN AV Equity
TRN IM Equity	ELE SM Equity
TRN IM Equity	ENEL IM Equity
TRN IM Equity	EDF FP Equity
TRN IM Equity	EOAN GR Equity
TRN IM Equity	EVN AV Equity
ELE SM Equity	ENEL IM Equity
ELE SM Equity	EDF FP Equity
ELE SM Equity	EOAN GR Equity
ELE SM Equity	EVN AV Equity
ENEL IM Equity	EDF FP Equity
ENEL IM Equity	EOAN GR Equity
ENEL IM Equity	EVN AV Equity
EDF FP Equity	EOAN GR Equity
EDF FP Equity	EVN AV Equity
EOAN GR Equity	EVN AV Equity

Table Q.13: EU stock pairs

ENG SM Equity	SRG IM Equity
ENG SM Equity	GAS SM Equity
SRG IM Equity	GAS SM Equity
A2A IM Equity	ACE IM Equity
A2A IM Equity	GSZ FP Equity
A2A IM Equity	HER IM Equity
A2A IM Equity	RWE GR Equity
A2A IM Equity	VIE FP Equity
ACE IM Equity	GSZ FP Equity
ACE IM Equity	HER IM Equity
ACE IM Equity	RWE GR Equity
ACE IM Equity	VIE FP Equity
GSZ FP Equity	HER IM Equity
GSZ FP Equity	RWE GR Equity
GSZ FP Equity	VIE FP Equity
HER IM Equity	RWE GR Equity
HER IM Equity	VIE FP Equity
RWE GR Equity	VIE FP Equity
EDN IM Equity	EEN FP Equity
EDN IM Equity	IBR SM Equity
EDN IM Equity	AGS SM Equity
EEN FP Equity	IBR SM Equity
EEN FP Equity	AGS SM Equity
IBR SM Equity	AGS SM Equity
FIE GR Equity	ITX SM Equity
FIE GR Equity	MEO GR Equity
FIE GR Equity	PP FP Equity
ITX SM Equity	MEO GR Equity
ITX SM Equity	PP FP Equity
MEO GR Equity	PP FP Equity

Table Q.14: EU stock pairs

G1A GR Equity	VK FP Equity
G1A GR Equity	ZOT SM Equity
G1A GR Equity	ANDR AV Equity
G1A GR Equity	SIE GR Equity
G1A GR Equity	MAN GR Equity
G1A GR Equity	ALO FP Equity
VK FP Equity	ZOT SM Equity
VK FP Equity	ANDR AV Equity
VK FP Equity	SIE GR Equity
VK FP Equity	MAN GR Equity
VK FP Equity	ALO FP Equity
ZOT SM Equity	ANDR AV Equity
ZOT SM Equity	SIE GR Equity
ZOT SM Equity	MAN GR Equity
ZOT SM Equity	ALO FP Equity
ANDR AV Equity	SIE GR Equity
ANDR AV Equity	MAN GR Equity
ANDR AV Equity	ALO FP Equity
SIE GR Equity	MAN GR Equity
SIE GR Equity	ALO FP Equity
MAN GR Equity	ALO FP Equity
BIM FP Equity	EI FP Equity
BIM FP Equity	FME GR Equity
BIM FP Equity	FRE GR Equity
BIM FP Equity	RHK GR Equity
EI FP Equity	FME GR Equity
EI FP Equity	FRE GR Equity
EI FP Equity	RHK GR Equity
FME GR Equity	FRE GR Equity
FME GR Equity	RHK GR Equity

TABLE Q.15: EU stock pairs

FRE GR Equity	RHK GR Equity
BB FP Equity	SK FP Equity
TKA GR Equity	ACX SM Equity
TKA GR Equity	MT NA Equity
TKA GR Equity	SZG GR Equity
TKA GR Equity	VOE AV Equity
TKA GR Equity	TEN IM Equity
TKA GR Equity	ERA FP Equity
ACX SM Equity	MT NA Equity
ACX SM Equity	SZG GR Equity
ACX SM Equity	VOE AV Equity
ACX SM Equity	TEN IM Equity
ACX SM Equity	ERA FP Equity
MT NA Equity	SZG GR Equity
MT NA Equity	VOE AV Equity
MT NA Equity	TEN IM Equity
MT NA Equity	ERA FP Equity
SZG GR Equity	VOE AV Equity
SZG GR Equity	TEN IM Equity
SZG GR Equity	ERA FP Equity
VOE AV Equity	TEN IM Equity
VOE AV Equity	ERA FP Equity
TEN IM Equity	ERA FP Equity
ARR FP Equity	ABE SM Equity
ARR FP Equity	ATL IM Equity
ARR FP Equity	FER SM Equity
ABE SM Equity	ATL IM Equity
ABE SM Equity	FER SM Equity
ATL IM Equity	FER SM Equity
DPW GR Equity	TNT NA Equity

Table Q.16: EU stock pairs

DPW GR Equity	POST AV Equity
TNT NA Equity	POST AV Equity
HHFA GR Equity	VPK NA Equity
ADP FP Equity	FRA GR Equity
PHIA NA Equity	UBI FP Equity
AGN NA Equity	CNP FP Equity
AGN NA Equity	MED IM Equity
CNP FP Equity	MED IM Equity
ALV GR Equity	CS FP Equity
ALV GR Equity	GCO SM Equity
ALV GR Equity	G IM Equity
ALV GR Equity	MAP SM Equity
CS FP Equity	GCO SM Equity
CS FP Equity	G IM Equity
CS FP Equity	MAP SM Equity
GCO SM Equity	G IM Equity
GCO SM Equity	MAP SM Equity
G IM Equity	MAP SM Equity
MUV2 GR Equity	SCR FP Equity
MUV2 GR Equity	HNR1 GR Equity
MUV2 GR Equity	VIG AV Equity
MUV2 GR Equity	ELE FP Equity
SCR FP Equity	HNR1 GR Equity
SCR FP Equity	VIG AV Equity
SCR FP Equity	ELE FP Equity
HNR1 GR Equity	VIG AV Equity
HNR1 GR Equity	ELE FP Equity
VIG AV Equity	ELE FP Equity
CEP SM Equity	ENI IM Equity
CEP SM Equity	OMV AV Equity

Table Q.17: EU stock pairs

CEP SM Equity	REP SM Equity
CEP SM Equity	FP FP Equity
ENI IM Equity	OMV AV Equity
ENI IM Equity	REP SM Equity
ENI IM Equity	FP FP Equity
OMV AV Equity	REP SM Equity
OMV AV Equity	FP FP Equity
REP SM Equity	FP FP Equity
ERG IM Equity	SRS IM Equity
ETL FP Equity	SKYD GR Equity
DEC FP Equity	PUB FP Equity
MMB FP Equity	PAJ FP Equity
MMB FP Equity	REN NA Equity
MMB FP Equity	WKL NA Equity
PAJ FP Equity	REN NA Equity
PAJ FP Equity	WKL NA Equity
REN NA Equity	WKL NA Equity
MMT FP Equity	MS IM Equity
MMT FP Equity	TL5 SM Equity
MMT FP Equity	TFI FP Equity
MS IM Equity	TL5 SM Equity
MS IM Equity	TFI FP Equity
TL5 SM Equity	TFI FP Equity
GBB FP Equity	GA FP Equity
GBB FP Equity	SPM IM Equity
GBB FP Equity	SBMO NA Equity
GBB FP Equity	TEC FP Equity
GA FP Equity	SPM IM Equity
GA FP Equity	SBMO NA Equity
GA FP Equity	TEC FP Equity

Table Q.18: EU stock pairs

SPM IM Equity	SBMO NA Equity
SPM IM Equity	TEC FP Equity
SBMO NA Equity	TEC FP Equity
ADS GR Equity	CDI FP Equity
ADS GR Equity	RMS FP Equity
ADS GR Equity	LUX IM Equity
ADS GR Equity	MC FP Equity
CDI FP Equity	RMS FP Equity
CDI FP Equity	LUX IM Equity
CDI FP Equity	MC FP Equity
RMS FP Equity	LUX IM Equity
RMS FP Equity	MC FP Equity
LUX IM Equity	MC FP Equity
BEI GR Equity	OR FP Equity
IPN FP Equity	MRK GR Equity
IPN FP Equity	SAN FP Equity
MRK GR Equity	SAN FP Equity
ATO FP Equity	CAP FP Equity
DSY FP Equity	SAP GR Equity
ILD FP Equity	SOW GR Equity
ILD FP Equity	UTDI GR Equity
SOW GR Equity	UTDI GR Equity
AF FP Equity	LHA GR Equity
AF FP Equity	IBLA SM Equity
LHA GR Equity	IBLA SM Equity
AGL IM Equity	SW FP Equity
FNC IM Equity	EAD FP Equity
SAF FP Equity	EAD FP Equity
HO FP Equity	EAD FP Equity
SAF FP Equity	FNC IM Equity

Table Q.19: EU stock pairs

HO FP Equity	FNC IM Equity
HO FP Equity	SAF FP Equity
DAI GR Equity	BMW GR Equity
F IM Equity	BMW GR Equity
UG FP Equity	BMW GR Equity
RNO FP Equity	BMW GR Equity
VOW GR Equity	BMW GR Equity
F IM Equity	DAI GR Equity
UG FP Equity	DAI GR Equity
RNO FP Equity	DAI GR Equity
VOW GR Equity	DAI GR Equity
UG FP Equity	F IM Equity
RNO FP Equity	F IM Equity
VOW GR Equity	F IM Equity
RNO FP Equity	UG FP Equity
VOW GR Equity	UG FP Equity
VOW GR Equity	RNO FP Equity
ML FP Equity	CON GR Equity
S92 GR Equity	GAM SM Equity
SWV GR Equity	GAM SM Equity
SWV GR Equity	S92 GR Equity
HEIA NA Equity	HEIO NA Equity
RI FP Equity	HEIO NA Equity
CPR IM Equity	HEIO NA Equity
RI FP Equity	HEIA NA Equity
CPR IM Equity	HEIA NA Equity
CPR IM Equity	RI FP Equity
BAS GR Equity	AKZA NA Equity
LXS GR Equity	AKZA NA Equity
WCH GR Equity	AKZA NA Equity

Table Q.20: EU stock pairs

DSM NA Equity	AKZA NA Equity
LXS GR Equity	BAS GR Equity
WCH GR Equity	BAS GR Equity
DSM NA Equity	BAS GR Equity
WCH GR Equity	LXS GR Equity
DSM NA Equity	LXS GR Equity
DSM NA Equity	WCH GR Equity
AI FP Equity	LIN GR Equity
BAYN GR Equity	LIN GR Equity
SDF GR Equity	LIN GR Equity
BAYN GR Equity	AI FP Equity
SDF GR Equity	AI FP Equity
SDF GR Equity	BAYN GR Equity
ACS SM Equity	ABG SM Equity
EN FP Equity	ABG SM Equity
FGR FP Equity	ABG SM Equity
FCC SM Equity	ABG SM Equity
HOT GR Equity	ABG SM Equity
SYV SM Equity	ABG SM Equity
STR AV Equity	ABG SM Equity
DG FP Equity	ABG SM Equity
EN FP Equity	ACS SM Equity
FGR FP Equity	ACS SM Equity
FCC SM Equity	ACS SM Equity
HOT GR Equity	ACS SM Equity
SYV SM Equity	ACS SM Equity
STR AV Equity	ACS SM Equity
DG FP Equity	ACS SM Equity
FGR FP Equity	EN FP Equity
FCC SM Equity	EN FP Equity

Table Q.21: EU stock pairs

HOT GR Equity	EN FP Equity
SYV SM Equity	EN FP Equity
STR AV Equity	EN FP Equity
DG FP Equity	EN FP Equity
FCC SM Equity	FGR FP Equity
HOT GR Equity	FGR FP Equity
SYV SM Equity	FGR FP Equity
STR AV Equity	FGR FP Equity
DG FP Equity	FGR FP Equity
HOT GR Equity	FCC SM Equity
SYV SM Equity	FCC SM Equity
STR AV Equity	FCC SM Equity
DG FP Equity	FCC SM Equity
SYV SM Equity	HOT GR Equity
STR AV Equity	HOT GR Equity
DG FP Equity	HOT GR Equity
STR AV Equity	SYV SM Equity
DG FP Equity	SYV SM Equity
DG FP Equity	STR AV Equity
HEI GR Equity	BZU IM Equity
IT IM Equity	BZU IM Equity
LG FP Equity	BZU IM Equity
IT IM Equity	HEI GR Equity
LG FP Equity	HEI GR Equity
LG FP Equity	IT IM Equity
TRE SM Equity	SGO FP Equity
ANA SM Equity	SGO FP Equity
ANA SM Equity	TRE SM Equity
PRY IM Equity	LR FP Equity
SU FP Equity	LR FP Equity

Table Q.22: EU stock pairs

SU FP Equity	PRY IM Equity
CRI SM Equity	ALB SM Equity
RF FP Equity	ALB SM Equity
EXO IM Equity	ALB SM Equity
RF FP Equity	CRI SM Equity
EXO IM Equity	CRI SM Equity
EXO IM Equity	RF FP Equity
DB1 GR Equity	BME SM Equity
KPN NA Equity	FTE FP Equity
TIT IM Equity	FTE FP Equity
TEF SM Equity	FTE FP Equity
TKA AV Equity	FTE FP Equity
TIT IM Equity	KPN NA Equity
TEF SM Equity	KPN NA Equity
TKA AV Equity	KPN NA Equity
TEF SM Equity	TIT IM Equity
TKA AV Equity	TIT IM Equity
TKA AV Equity	TEF SM Equity
CLS1 GR Equity	CA FP Equity
CO FP Equity	AH NA Equity
BMPS IM Equity	CRG IM Equity
PMI IM Equity	CRG IM Equity
BP IM Equity	CRG IM Equity
POP SM Equity	CRG IM Equity
SAB SM Equity	CRG IM Equity
SAN SM Equity	CRG IM Equity
BKT SM Equity	CRG IM Equity
BBVA SM Equity	CRG IM Equity
BNP FP Equity	CRG IM Equity
CBK GR Equity	CRG IM Equity

Table Q.23: EU stock pairs

ACA FP Equity	CRG IM Equity
DBK GR Equity	CRG IM Equity
EUROB GA Equity	CRG IM Equity
EBS AV Equity	CRG IM Equity
ISP IM Equity	CRG IM Equity
MB IM Equity	CRG IM Equity
KN FP Equity	CRG IM Equity
GLE FP Equity	CRG IM Equity
UBI IM Equity	CRG IM Equity
UCG IM Equity	CRG IM Equity
BTO SM Equity	CRG IM Equity
DPB GR Equity	CRG IM Equity
PMI IM Equity	BMPS IM Equity
BP IM Equity	BMPS IM Equity
POP SM Equity	BMPS IM Equity
SAB SM Equity	BMPS IM Equity
SAN SM Equity	BMPS IM Equity
BKT SM Equity	BMPS IM Equity
BBVA SM Equity	BMPS IM Equity
BNP FP Equity	BMPS IM Equity
CBK GR Equity	BMPS IM Equity
ACA FP Equity	BMPS IM Equity
DBK GR Equity	BMPS IM Equity
EUROB GA Equity	BMPS IM Equity
EBS AV Equity	BMPS IM Equity
ISP IM Equity	BMPS IM Equity
MB IM Equity	BMPS IM Equity
KN FP Equity	BMPS IM Equity
GLE FP Equity	BMPS IM Equity
UBI IM Equity	BMPS IM Equity

Table Q.24: EU stock pairs

UCG IM Equity	BMPS IM Equity
BTO SM Equity	BMPS IM Equity
DPB GR Equity	BMPS IM Equity
BP IM Equity	PMI IM Equity
POP SM Equity	PMI IM Equity
SAB SM Equity	PMI IM Equity
SAN SM Equity	PMI IM Equity
BKT SM Equity	PMI IM Equity
BBVA SM Equity	PMI IM Equity
BNP FP Equity	PMI IM Equity
CBK GR Equity	PMI IM Equity
ACA FP Equity	PMI IM Equity
DBK GR Equity	PMI IM Equity
EUROB GA Equity	PMI IM Equity
EBS AV Equity	PMI IM Equity
ISP IM Equity	PMI IM Equity
MB IM Equity	PMI IM Equity
KN FP Equity	PMI IM Equity
GLE FP Equity	PMI IM Equity
UBI IM Equity	PMI IM Equity
UCG IM Equity	PMI IM Equity
BTO SM Equity	PMI IM Equity
DPB GR Equity	PMI IM Equity
POP SM Equity	BP IM Equity
SAB SM Equity	BP IM Equity
SAN SM Equity	BP IM Equity
BKT SM Equity	BP IM Equity
BBVA SM Equity	BP IM Equity
BNP FP Equity	BP IM Equity
CBK GR Equity	BP IM Equity

Table Q.25: EU stock pairs

ACA FP Equity	BP IM Equity
DBK GR Equity	BP IM Equity
EUROB GA Equity	BP IM Equity
EBS AV Equity	BP IM Equity
ISP IM Equity	BP IM Equity
MB IM Equity	BP IM Equity
KN FP Equity	BP IM Equity
GLE FP Equity	BP IM Equity
UBI IM Equity	BP IM Equity
UCG IM Equity	BP IM Equity
BTO SM Equity	BP IM Equity
DPB GR Equity	BP IM Equity
SAB SM Equity	POP SM Equity
SAN SM Equity	POP SM Equity
BKT SM Equity	POP SM Equity
BBVA SM Equity	POP SM Equity
BNP FP Equity	POP SM Equity
CBK GR Equity	POP SM Equity
ACA FP Equity	POP SM Equity
DBK GR Equity	POP SM Equity
EUROB GA Equity	POP SM Equity
EBS AV Equity	POP SM Equity
ISP IM Equity	POP SM Equity
MB IM Equity	POP SM Equity
KN FP Equity	POP SM Equity
GLE FP Equity	POP SM Equity
UBI IM Equity	POP SM Equity
UCG IM Equity	POP SM Equity
BTO SM Equity	POP SM Equity
DPB GR Equity	POP SM Equity

Table Q.26: EU stock pairs

SAN SM Equity	SAB SM Equity
BKT SM Equity	SAB SM Equity
BBVA SM Equity	SAB SM Equity
BNP FP Equity	SAB SM Equity
CBK GR Equity	SAB SM Equity
ACA FP Equity	SAB SM Equity
DBK GR Equity	SAB SM Equity
EUROB GA Equity	SAB SM Equity
EBS AV Equity	SAB SM Equity
ISP IM Equity	SAB SM Equity
MB IM Equity	SAB SM Equity
KN FP Equity	SAB SM Equity
GLE FP Equity	SAB SM Equity
UBI IM Equity	SAB SM Equity
UCG IM Equity	SAB SM Equity
BTO SM Equity	SAB SM Equity
DPB GR Equity	SAB SM Equity
BKT SM Equity	SAN SM Equity
BBVA SM Equity	SAN SM Equity
BNP FP Equity	SAN SM Equity
CBK GR Equity	SAN SM Equity
ACA FP Equity	SAN SM Equity
DBK GR Equity	SAN SM Equity
EUROB GA Equity	SAN SM Equity
EBS AV Equity	SAN SM Equity
ISP IM Equity	SAN SM Equity
MB IM Equity	SAN SM Equity
KN FP Equity	SAN SM Equity
GLE FP Equity	SAN SM Equity
UBI IM Equity	SAN SM Equity

Table Q.27: EU stock pairs

UCG IM Equity	SAN SM Equity
BTO SM Equity	SAN SM Equity
DPB GR Equity	SAN SM Equity
BTO SM Equity	DBK GR Equity
DPB GR Equity	DBK GR Equity
EBS AV Equity	EUROB GA Equity
ISP IM Equity	EUROB GA Equity
KBC BB Equity	EUROB GA Equity
MB IM Equity	EUROB GA Equity
KN FP Equity	EUROB GA Equity
GLE FP Equity	EUROB GA Equity
UBI IM Equity	EUROB GA Equity
UCG IM Equity	EUROB GA Equity
BTO SM Equity	EUROB GA Equity
DPB GR Equity	EUROB GA Equity
ISP IM Equity	EBS AV Equity
MB IM Equity	EBS AV Equity
KN FP Equity	EBS AV Equity
GLE FP Equity	EBS AV Equity
UBI IM Equity	EBS AV Equity
UCG IM Equity	EBS AV Equity
BTO SM Equity	EBS AV Equity
DPB GR Equity	EBS AV Equity
MB IM Equity	ISP IM Equity
KN FP Equity	ISP IM Equity
GLE FP Equity	ISP IM Equity
UBI IM Equity	ISP IM Equity
UCG IM Equity	ISP IM Equity
BTO SM Equity	ISP IM Equity
DPB GR Equity	ISP IM Equity

Table Q.28: EU stock pairs

KN FP Equity	MB IM Equity
GLE FP Equity	MB IM Equity
UBI IM Equity	MB IM Equity
UCG IM Equity	MB IM Equity
BTO SM Equity	MB IM Equity
DPB GR Equity	MB IM Equity
GLE FP Equity	KN FP Equity
UBI IM Equity	KN FP Equity
UCG IM Equity	KN FP Equity
BTO SM Equity	KN FP Equity
DPB GR Equity	KN FP Equity
UBI IM Equity	GLE FP Equity
UCG IM Equity	GLE FP Equity
BTO SM Equity	GLE FP Equity
DPB GR Equity	GLE FP Equity
UCG IM Equity	UBI IM Equity
BTO SM Equity	UBI IM Equity
DPB GR Equity	UBI IM Equity
BTO SM Equity	UCG IM Equity
DPB GR Equity	UCG IM Equity
DPB GR Equity	BTO SM Equity
EVA SM Equity	BN FP Equity
PLT IM Equity	BN FP Equity
PLT IM Equity	EVA SM Equity
VER AV Equity	IBE SM Equity
REE SM Equity	IBE SM Equity
TRN IM Equity	IBE SM Equity
ELE SM Equity	IBE SM Equity
ENEL IM Equity	IBE SM Equity
EDF FP Equity	IBE SM Equity

Table Q.29: EU stock pairs

EOAN GR Equity	IBE SM Equity
EVN AV Equity	IBE SM Equity
REE SM Equity	VER AV Equity
TRN IM Equity	VER AV Equity
ELE SM Equity	VER AV Equity
ENEL IM Equity	VER AV Equity
EDF FP Equity	VER AV Equity
EOAN GR Equity	VER AV Equity
EVN AV Equity	VER AV Equity
TRN IM Equity	REE SM Equity
ELE SM Equity	REE SM Equity
ENEL IM Equity	REE SM Equity
EDF FP Equity	REE SM Equity
EOAN GR Equity	REE SM Equity
EVN AV Equity	REE SM Equity
ELE SM Equity	TRN IM Equity
ENEL IM Equity	TRN IM Equity
EDF FP Equity	TRN IM Equity
EOAN GR Equity	TRN IM Equity
EVN AV Equity	TRN IM Equity
ENEL IM Equity	ELE SM Equity
EDF FP Equity	ELE SM Equity
EOAN GR Equity	ELE SM Equity
EVN AV Equity	ELE SM Equity
EDF FP Equity	ENEL IM Equity
EOAN GR Equity	ENEL IM Equity
EVN AV Equity	ENEL IM Equity
EOAN GR Equity	EDF FP Equity
EVN AV Equity	EDF FP Equity
EVN AV Equity	EOAN GR Equity

Table Q.30: EU stock pairs

SRG IM Equity	ENG SM Equity
GAS SM Equity	ENG SM Equity
GAS SM Equity	SRG IM Equity
ACE IM Equity	A2A IM Equity
GSZ FP Equity	A2A IM Equity
HER IM Equity	A2A IM Equity
RWE GR Equity	A2A IM Equity
VIE FP Equity	A2A IM Equity
GSZ FP Equity	ACE IM Equity
HER IM Equity	ACE IM Equity
RWE GR Equity	ACE IM Equity
VIE FP Equity	ACE IM Equity
HER IM Equity	GSZ FP Equity
RWE GR Equity	GSZ FP Equity
VIE FP Equity	GSZ FP Equity
RWE GR Equity	HER IM Equity
VIE FP Equity	HER IM Equity
VIE FP Equity	RWE GR Equity
EEN FP Equity	EDN IM Equity
IBR SM Equity	EDN IM Equity
AGS SM Equity	EDN IM Equity
IBR SM Equity	EEN FP Equity
AGS SM Equity	EEN FP Equity
AGS SM Equity	IBR SM Equity
ITX SM Equity	FIE GR Equity
MEO GR Equity	FIE GR Equity
PP FP Equity	FIE GR Equity
MEO GR Equity	ITX SM Equity
PP FP Equity	ITX SM Equity
PP FP Equity	MEO GR Equity

Table Q.31: EU stock pairs

VK FP Equity	G1A GR Equity
ZOT SM Equity	G1A GR Equity
ANDR AV Equity	G1A GR Equity
SIE GR Equity	G1A GR Equity
MAN GR Equity	G1A GR Equity
ALO FP Equity	G1A GR Equity
ZOT SM Equity	VK FP Equity
ANDR AV Equity	VK FP Equity
SIE GR Equity	VK FP Equity
MAN GR Equity	VK FP Equity
ALO FP Equity	VK FP Equity
ANDR AV Equity	ZOT SM Equity
SIE GR Equity	ZOT SM Equity
MAN GR Equity	ZOT SM Equity
ALO FP Equity	ZOT SM Equity
SIE GR Equity	ANDR AV Equity
MAN GR Equity	ANDR AV Equity
ALO FP Equity	ANDR AV Equity
MAN GR Equity	SIE GR Equity
ALO FP Equity	SIE GR Equity
ALO FP Equity	MAN GR Equity
EI FP Equity	BIM FP Equity
FME GR Equity	BIM FP Equity
FRE GR Equity	BIM FP Equity
RHK GR Equity	BIM FP Equity
FME GR Equity	EI FP Equity
FRE GR Equity	EI FP Equity
RHK GR Equity	EI FP Equity
FRE GR Equity	FME GR Equity
RHK GR Equity	FME GR Equity

Table Q.32: EU stock pairs

RHK GR Equity	FRE GR Equity
SK FP Equity	BB FP Equity
ACX SM Equity	TKA GR Equity
MT NA Equity	TKA GR Equity
SZG GR Equity	TKA GR Equity
VOE AV Equity	TKA GR Equity
TEN IM Equity	TKA GR Equity
ERA FP Equity	TKA GR Equity
MT NA Equity	ACX SM Equity
SZG GR Equity	ACX SM Equity
VOE AV Equity	ACX SM Equity
TEN IM Equity	ACX SM Equity
ERA FP Equity	ACX SM Equity
SZG GR Equity	MT NA Equity
VOE AV Equity	MT NA Equity
TEN IM Equity	MT NA Equity
ERA FP Equity	MT NA Equity
VOE AV Equity	SZG GR Equity
TEN IM Equity	SZG GR Equity
ERA FP Equity	SZG GR Equity
TEN IM Equity	VOE AV Equity
ERA FP Equity	VOE AV Equity
ERA FP Equity	TEN IM Equity
ABE SM Equity	ARR FP Equity
ATL IM Equity	ARR FP Equity
FER SM Equity	ARR FP Equity
ATL IM Equity	ABE SM Equity
FER SM Equity	ABE SM Equity
FER SM Equity	ATL IM Equity
TNT NA Equity	DPW GR Equity

Table Q.33: EU stock pairs

POST AV Equity	DPW GR Equity
POST AV Equity	TNT NA Equity
VPK NA Equity	HHFA GR Equity
FRA GR Equity	ADP FP Equity
UBI FP Equity	PHIA NA Equity
CNP FP Equity	AGN NA Equity
MED IM Equity	AGN NA Equity
MED IM Equity	CNP FP Equity
CS FP Equity	ALV GR Equity
GCO SM Equity	ALV GR Equity
G IM Equity	ALV GR Equity
MAP SM Equity	ALV GR Equity
GCO SM Equity	CS FP Equity
G IM Equity	CS FP Equity
MAP SM Equity	CS FP Equity
G IM Equity	GCO SM Equity
MAP SM Equity	GCO SM Equity
MAP SM Equity	G IM Equity
SCR FP Equity	MUV2 GR Equity
HNR1 GR Equity	MUV2 GR Equity
VIG AV Equity	MUV2 GR Equity
ELE FP Equity	MUV2 GR Equity
HNR1 GR Equity	SCR FP Equity
VIG AV Equity	SCR FP Equity
ELE FP Equity	SCR FP Equity
VIG AV Equity	HNR1 GR Equity
ELE FP Equity	HNR1 GR Equity
ELE FP Equity	VIG AV Equity
ENI IM Equity	CEP SM Equity
OMV AV Equity	CEP SM Equity

Table Q.34: EU stock pairs

REP SM Equity	CEP SM Equity
FP FP Equity	CEP SM Equity
OMV AV Equity	ENI IM Equity
REP SM Equity	ENI IM Equity
FP FP Equity	ENI IM Equity
REP SM Equity	OMV AV Equity
FP FP Equity	OMV AV Equity
FP FP Equity	REP SM Equity
SRS IM Equity	ERG IM Equity
SKYD GR Equity	ETL FP Equity
PUB FP Equity	DEC FP Equity
PAJ FP Equity	MMB FP Equity
REN NA Equity	MMB FP Equity
WKL NA Equity	MMB FP Equity
REN NA Equity	PAJ FP Equity
WKL NA Equity	PAJ FP Equity
WKL NA Equity	REN NA Equity
MS IM Equity	MMT FP Equity
TL5 SM Equity	MMT FP Equity
TFI FP Equity	MMT FP Equity
TL5 SM Equity	MS IM Equity
TFI FP Equity	MS IM Equity
TFI FP Equity	TL5 SM Equity
GA FP Equity	GBB FP Equity
SPM IM Equity	GBB FP Equity
SBMO NA Equity	GBB FP Equity
TEC FP Equity	GBB FP Equity
SPM IM Equity	GA FP Equity
SBMO NA Equity	GA FP Equity
TEC FP Equity	GA FP Equity

Table Q.35: EU stock pairs

SBMO NA Equity	SPM IM Equity
TEC FP Equity	SPM IM Equity
TEC FP Equity	SBMO NA Equity
CDI FP Equity	ADS GR Equity
RMS FP Equity	ADS GR Equity
LUX IM Equity	ADS GR Equity
MC FP Equity	ADS GR Equity
RMS FP Equity	CDI FP Equity
LUX IM Equity	CDI FP Equity
MC FP Equity	CDI FP Equity
LUX IM Equity	RMS FP Equity
MC FP Equity	RMS FP Equity
MC FP Equity	LUX IM Equity
OR FP Equity	BEI GR Equity
MRK GR Equity	IPN FP Equity
SAN FP Equity	IPN FP Equity
SAN FP Equity	MRK GR Equity
CAP FP Equity	ATO FP Equity
IDR SM Equity	CAP FP Equity
SAP GR Equity	DSY FP Equity
SOW GR Equity	ILD FP Equity
UTDI GR Equity	ILD FP Equity
UTDI GR Equity	SOW GR Equity
LHA GR Equity	AF FP Equity
IBLA SM Equity	AF FP Equity
IBLA SM Equity	LHA GR Equity
SW FP Equity	AGL IM Equity

References

J. Li Y. Ait-Sahalia, U. Fan. The leverage effect puzzle: Disentangling sources of bias at high frequency., November 2011. URL: http://www.nber.org/papers/w17592.pdf.

M. A. Al Halaseh, R. H. S. Islam and R. Bakar. Dynamic portfolio selection: A literature revisit. *International Business Management*, 10:67–77, 2016.

N. Amenc and L Martellini. Portfolio optimization and hedge fund style allocation decisions. *The Journal of Alternative Investments*, Vol. 5, No. 2:7–20, 2002.

A. Pedersen L Asness, C. S. Franzzini. Leverage aversion and risk parity. *Journal of financial analysts*, 68, 2012.

C.S. Asness, T. J. Moskowitz, and L.H. Pedersen. Value and momentum everywhere. *The Journal of Finance*, 68:929–985, 2013.

S C. Asness. The interaction of value and momentum strategies. *Financial Analysts Journal*, 53:29–36, 1997.

Y. Balvers, R. J. & Wu. Momentum and mean reversion across national equity markets. *Journal of Emprical Finance*, 13:24–48, 2006.

J. Banks, R. Blundell, R. Disney, and C. Ennerson. Retirement, pensions and the adequacy of saving: A guide to the debate. Published online, http://www.ifs.org.uk/bns/bn29.pdf, October 2002. Briefing Note.

F. Black, M. Jensen, and M. Scholes. The capital asset pricing model: Some empirical tests. *Studies in the Theory of Capital Markets, Praeger: New York*, pages 79–124, 1972.

F. Black and R. Litterman. Global portfolio optimization. *Financial Analysts Journal*, September/October:28–43, 1992.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

A. Blum and A Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35:193–205, 1999.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

T. Bollerslev. Modeling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *The Review of Economics and Statistics*, 72:498–505, 1990.

G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control (Wiley Series in Probability and Statistics)*. Wiley & Co., 2008.

Bradley. *Distribution-Free Statistical Tests*. Prentice-Hall, 1968.

L. Breiman. Optimal gambling systems for favorable games. *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1:65–78, 1961.

A. N. Burgess. *A Computational Methodology for Modelling the Dynamics of Statistical Arbitrage*. PhD thesis, University of London, 1999.

Y. Cai, K. Judd, and R. Xu. Numerical solution of dynamic portfolio optimization with transaction costs. Electronic, January 2013. URL: http://econpapers.repec.org/paper/nbrnberwo/18709.htm.

V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19:6–11, 1993.

N. A. Chriss and R. Almgren. Portfolios from sorts. *Working paper*, 2005.

R. Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance Volume 1 (2001) 223236*, 1:223–236, 2001.

M. Covel. *Trend Following: How Great Traders Make Millions in Up or Down Markets*. FT Press, 2007.

T. M. Cover. Universal portfolios. *Mathematical Finance*, 1:1–29, 1991.

T.M. Cover and E. Ordentlich. Universal portfolios with side information. *Information Theory, IEEE Transactions*, 42:348–363, 1996.

F. Cross. The behavior of stock prices on Fridays and Mondays,. *Financial Analysts Journal*, 29:67–69, 1973.

M. Davis and S. Lleo. *Fractional Kelly Strategies in Continious Time Recent Developments*. Handbook of Fundamentals of Financial Decision making. World Scientific Publishing, 2012.

R. De Bondt, W. F. M. & Thaler. Does the stock market overreact? *The Journal of Finance*, 40(3):793805, 1985.

L. & Uppal R. DeMiguel, V. Garlappi. How inefficient is the 1/n asset allocation strategy?, 2009. URL: http://faculty.london.edu/avmiguel/DeMiguel-Garlappi-Uppal-RFS.pdf.

R.J. Elliott, J. V .D. Hoek, and W. P. Malcolm. Pairs trading. *Quantitative Finance*, 5 No.3:271–276, 2005.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica*, 50:987–1008, 1982.

E. Fama and K. French. Permanent and temporary components of stock prices. *Journal of Political Economy*, 96:246–273, 1988.

E. Fama and K. French. The capital asset pricing model: Theory and evidence. *The Journal of Economic Perspectives*, 18:25–46, 2004.

A. Frazzini and L. H. Pedersen. Betting againt beta. *Journal of Financial Economics*, 111:1–25, 2014.

K. R. French. Stock returns and the weekend effect. *Journal of Financial Economics*, 8:55–69, 1980.

A. & Wunderlich R. Frey, R. & Gabih. Portfolio optimization under partial information with expert opinions. *International Journal of Theoretical and Applied Finance*, 15: 1–18, 2012.

E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst. Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19:797–827, 2006.

A Ghosh and A. Mahanti. Investment portfolio management: A review from 2009 to 2014. In *Proceedings of 10th Global Business and Social Science Research Conference, 23 -24 June 2014, Beijing, China.*, 2014.

R Grinold and R. Kahn. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. McGraw-Hill, 1999.

L Gyorfi, F. Udina, and H. Walk. Experiments on universal portfolio selection using data from real markets. *Working paper*, 2008.

R. A. Haugen and J. Lakonishok. *The Incredible January Effect: The Stock Markets Unsolved Mystery*. Dow Jones-Irwin, Homewood, IL., 1987.

T. Henker, J. Henker and T. D. Huynh. Survivorship bias and alternative explanations of momentum effect. 23rd Australasian Finance and Banking Conference 2010 Paper. Electronic, 2010. URL: http://doi.org/10.2139/ssrn.1663495.

D. A. & Chicken E. Hollander, M. & Wolgfe. *Nonparametric Statistical Methods*. John Wiley & Co., 2014.

H. Hong and J. Stein. A unified theory of underreaction, momentum trading and over reaction in asset markets. *Journal of Finance*, 54:2143–2184, 1999.

H. S. Houthakker. Systematic and random elements in short-term price movements. *The American Economic Review*, 51:164–172, 1961.

J. Hull. *Options Futures and Other Derivatives*. Pearson Education, 2011.

A. Javaheri. *Inside Volatility Arbitrage: The Secrets of Skewness*. John Wiley & Sons, Inc., 2005.

N. Jegadeesh and S. Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48:65–91, 1993.

N. Jegadeesh and S. Titman. Profitability of momentum strategies: An evaluation of alternative explainations. *Journal of Finance*, 56:699–720, 2001.

M. C. Jensen. The performance of mutual funds in the period 1945 - 1964. *The Journal of Finance*, 23:389–416, 1968.

M. C. Jensen, F. Black, and M. S. Scholes. *The Capital Asset Pricing Model: Some Empirical Tests*. Praeger, 1972.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82:35–45, 1960.

J. Kelly. A new interpretation of information rate. *Information Theory*, 2:185–189, 1956.

M. G. Kendall. The analysis of economic time-series - part I: Prices. *Journal of the Royal Statistical Society*, 116:11–25, 1953.

Phil Kim. *Kalman Filter for Beginners with MATLAB Examples*. A-JIN Publishing company, 2011.

S. S. Kozat and C.A. Singer. Universal semiconstant rebalanced portfolios. *Mathematical Finance*, 21:293–311, 2011.

P. Laureti, M. Medo, and Y. C. Zhang. Analysis of kelly-optimal portfolios. *Journal of Quantitative Finance*, 10 issue 7:689–97, 2010.

C. Lee and B. Swaminathan. Price momentum and trading volume. *Journal of Finance*, 55:2017–2070, 2000.

Y. Lemperiere, C. Deremble, P. Seager, M. Potters, and J.P. Bouchaud. Two centuries of trend following. Electronic, April 2014. URL: http://arxiv.org/abs/1404.3274.

J. Lewellen. Momentum and autocorrelation in stock returns. *Review of Financial Studies*, 15(2):533–564, 2002.

F. L'habitant. *Handbook of Hedge Funds*. John Wiley & Sons, Inc., 2007.

J. Lintner. The valuation of risk assets and selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1):13–37, 1965.

L. Liu and L. Zhang. Momentum profits, factor pricing, and macro-economic risk. *Review of Financial Studies*, 21:2417–2448, 2008.

A. W. Lo and A. C. MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, 1:41–66, 1988.

Donald Mackenzie. *An Engine, Not a Camera: How Financial Models Shape Markets.* MIT Press, 2008.

L.C. MacLean, E. O. Thorp, and W. T. Ziemba, editors. *The Kelly Capital Growth Investment Criterion: Theory and Practice.* World Scientific Press, 2011.

N. Mahler. *Modeling the S&P 500 Index Using the Kalman Filter and the LagLasso.* Machine Learning for Signal Processing, 2009. MLSP 2009. IEEE International Workshop, 2009.

T. Teiletche J. Maillard, S. Roncalli. On the properties of equally-weighted risk contributions portfolios. *Journal of Portfolio Management*, 36:60–70, 2010.

H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

H Markowitz. *Portfolio Selection Effecient Diversification of Investments.* John Wiley & Sons, Inc., New York, Chapman & Hall, Limited, London, 1959.

S. Maslov and Y. C. Zhang. Optimal investment strategy for risky assets. *International Journal of Theoretical and Applied Finance*, 1:377–387, 1998.

C. S. Mehra, A. Prugel-Bennett, E. Gerding, and V. Robu. Constructing smart portfolios from data driven quantitative investment models. In *IEEE Computational Intelligence for Financial Engineering and Economics, London, UK*, 2014.

R. C. Merton. An analytic derivation of the efficient frontier. *The Journal of Financial and Quantitative Analysis*, 7:1851–1872, 1972.

R. C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41:867–887, 1973.

R. C. Merton. *Continuous-Time Finance.* Wiley and Co, 1993.

R. O. Michaud. The markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal*, 45:31–42, 1989.

J. T. Moskowitz, Y. H. Ooib, and L.H. Pedersenb. Time series momentum. *Journal of Financial Economics*, 104:228–250, 2011.

M. F. M. Osborne. Periodic structure in the brownian motion of stock prices. *Operations Research*, pages 345–379, 1962.

J. K. Rising and A. J. Wyner. Partial kelly portfolios and shrinkage estimators. *Information Theory Proceedings (ISIT), IEEE International Symposium*, pages 1618–1622, 2012.

R. Ross, S. Roll. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13:341–360, 1976.

R. Ross, S. Roll. An empirical investigation of the arbitrage pricing theory. *The Journal of Finance*, 35:1073–1103, 1980.

A. F. Serban. Combining mean reversion and momentum trading strategies in foreign exchange markets. *Journal of Banking and Finance*, 34:2720–2727, 2010.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, pages 425–442, 1964.

E. Sinclair. *Volatility Trading*. Wiley and Co, 2008.

L. H. Summers. Does the stock market rationally reflect fundamental values? *Journal of Finance*, 41:591–601, 1986.

E. O. Thorp. *Beat the Market: A Scientific Stock Market System*. Random House, 1967.

A Timmermann, D. Blake, and Bruce N. Lehmann. Asset allocation dynamics and pension fund performance. *The Journal of Business*, 72:459–461, 1999.

J. Tobin. Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25:65–86, 1958.

H. Valian. *Optimisation Dynamic Portfolio Selection*. PhD thesis, Rutgers University, 2009.

G. Vidyamurthy. *Pairs Trading: Quantitative Methods and Analysis*. Wiley & Co., 2004.

S. B. Wachtel. Certain observations on seasonal movements in stock prices. *The Journal of Business of the University of Chicago*, pages 184–193, 1942.

C Wells. *The Kalman Filter in Finance*. Springer, 1996.

Paul Wilmott. *Paul Wilmott on Quantitative Finance*. John Wiley & Sons, 2006.