Portfolio Optimisation Models

Cristiano Arbex Valle

A thesis submitted for the degree of Doctor of Philosophy

School of Information Systems, Computing and Mathematics,
Brunel University

2013

Abstract

In this thesis we consider three different problems in the domain of portfolio optimisation. The first problem we consider is that of selecting an Absolute Return Portfolio (ARP). ARPs are usually seen as financial portfolios that aim to produce a good return regardless of how the underlying market performs, but our literature review shows that there is little agreement on what constitutes an ARP. We present a clear definition via a three-stage mixed-integer zero-one program for the problem of selecting an ARP.

The second problem considered is that of designing a Market Neutral Portfolio (MNP). MNPs are generally defined as financial portfolios that (ideally) exhibit performance independent from that of an underlying market, but, once again, the existing literature is very fragmented. We consider the problem of constructing a MNP as a mixed-integer non-linear program (MINLP) which minimises the absolute value of the correlation between portfolio return and underlying benchmark return.

The third problem is related to Exchange-Traded Funds (ETFs). ETFs are funds traded on the open market which typically have their performance tied to a benchmark index. They are composed of a basket of assets; most attempt to reproduce the returns of an index, but a growing number try to achieve a multiple of the benchmark return, such as two times or the negative of the return. We present a detailed performance study of the current ETF market and we find, among other conclusions, constant underperformance among ETFs that aim to do more than simply track an index. We present a MINLP for the problem of selecting the basket of assets that compose an ETF, which, to the best of our knowledge, is the first in the literature.

For all three models we present extensive computational results for portfolios derived from universes defined by S&P international equity indices with up to 1200 stocks. We use CPLEX to solve the ARP problem and the software package Minotaur for both our MINLPs for MNP and an ETF.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Ever since the pioneering work of Markowitz (1952), optimisation has been at the centre of work concerned with deciding the composition of financial portfolios. As such, both practitioners and academic researchers have been willing to trade off the disadvantages of optimisation (multiple optimal solutions, solution sensitivity) for its advantages (clear modelling framework, computational efficiency, algorithmic decision making).

The most popular portfolio optimisation problem is that of minimising risk for a given target expected return, or, conversely, maximising expected return while constraining risk. Different approaches measure risk differently, examples of different risk measures are variance of returns, CVar (Conditional Value at Risk, Rockafellar & Uryasev (2000)) and Sortino ratio (Sortino & van der Meer (1991)). Another popular portfolio optimisation problem is that of index tracking, where the concern is to reproduce (track) the performance of a financial index. Once again, there are different ways to measure tracking performance and hence there are different (and sometimes non-comparable) models for this purpose.

In short, different types of portfolios require different mathematical models, and, even for portfolios intended for the same purpose, the model to use is not uniquely defined. In this thesis we examine three portfolio optimisation problems that are not clearly defined in the present literature. We introduce optimisation models for the problems of selecting an Absolute Return Portfolio (ARP), a Market Neutral Portfolio (MNP) and the basket underlying an Exchange-Traded Fund (ETF).

ARPs are generally defined as financial portfolios that aim to produce good returns regardless of how the underlying market performs. However, our literature review shows that there is barely any agreement on what exactly defines a portfolio as an ARP. MNPs are defined as financial portfolios that (ideally) exhibit performance independent from that of an underlying market. Once again, there are numerous models and different ways to measure the independence of a MNP. Based on these definitions, both problems seem similar and, indeed, one can think of a portfolio that is independent from a benchmark as one which produces absolute returns irrespective of how the benchmark is performing. These problems are however different, and, in this thesis, we present clear and unique definitions and we introduce optimisation models for both of them.

ETFs, on the other hand, are not *per se* a portfolio model. They are funds that are traded on the open market and that usually have their expected performance tied to a benchmark index. ETFs are composed of a basket of assets held by the ETF creator and shares that are issued and traded in the open market. An ETF share entitles its holder to a portion of the underlying basket. Most ETFs are index trackers, however, some seek to achieve a multiple of benchmark return. For example, some ETFs aim to achieve the negative of index return, while others seek to achieve twice index return. The former are called inverse ETFs and the latter are known as leveraged ETFs. In this thesis we also examine the optimisation problem of defining an ETF basket of assets, a non-trivial problem especially for leveraged and inverse ETFs.

1.2 Thesis outline

The structure of this thesis is as follows. In Chapter 2 we present a literature survey of portfolio optimisation in general with special attention to ARPs, MNPs and ETFs. We start by summarising the history and context of portfolio optimisation theory. We then dedicate separate sections to the literature on each of the three main problems studied in this thesis.

In Chapter 3 we consider the problem of selecting an ARP. We present a three-stage mixed-integer zero-one program for the problem that explicitly considers transaction costs associated with trading. The first two stages relate to a regression of portfolio return against time, whilst the third stage relates to minimising transaction costs. We extend our approach to the problem of designing portfolios with differing characteristics. Computational results are given for portfolios of eleven different problem instances derived from universes defined by S&P international equity indices.

In Chapter 4 we consider the problem of constructing a MNP. We formulate this problem as a mixed-integer nonlinear program (MINLP), minimising the absolute value of the correlation between portfolio return and index return. Our model is a flexible one that incorporates decisions as to both long and short positions in assets. Computational results, obtained using the software package Minotaur, are given for the same problem instances as in Chapter 4. We also compare our approach against an alternative approach based on minimising the absolute value of regression slope (the zero-beta approach).

In Chapter 5 we present a survey of the current ETF market by collecting and analysing a large snapshot of 8192 ETFs, which compose the vast majority of the ETF market. We selected a subset of 822 ETFs to analyse more fully. Our performance analysis covers the period from January 1993 to September 2011 and statistically analyses these 822 ETFs, which have a total market value of US$1.81 trillion, using over 1.1m daily return observations. The accuracy with which ETFs replicate the behaviour of their benchmark is a mixed story; only 19% of ETFs reproduce both the mean return and the volatility of their benchmark within 1% p.a.. With respect to replicating benchmark volatility we found that most ETFs have higher volatility than their benchmarks.

Following the ETF survey performed in Chapter 5, we consider in Chapter 6 the problem of deciding the portfolio of assets that should underlie an ETF. We formulate this problem as a MINLP. We mostly consider ETFs which have positive leverage with respect to their benchmark index, as opposed to ETFs which simply attempt to track the benchmark performance, and ETFs which have negative leverage (inverse ETFs). Our formulation is a flexible one that incorporates decisions as to both long and short positions in assets, as well as including rebalancing and transaction costs. Computational results are given for problems for the same set of instances as used in Chapter 3. We also computationally compare our model to a previous model in the literature for index tracking.

Finally, in Chapter 7 we summarise the main results of our research, highlighting the contribution to knowledge we have made, and suggest directions for future work.

Chapter 2

Literature review

In this chapter we present a brief historical overview of portfolio optimisation and discuss studies in the literature related to Absolute Return and Market Neutral portfolios. We also discuss work related to Exchange-Traded Funds.

2.1 History of portfolio theory

The foundations of Modern Portfolio Theory (MPT) date back to the 1950s thanks to a landmark article and subsequent book by Markowitz (1952, 1959). Prior to his work, assets were analysed individually in order to construct a portfolio. Markowitz proposed that portfolios should be selected based on overall (instead of individual) risk-return assessment. An important assumption of MPT is that investors are risk averse, meaning that given two portfolios that offer the same expected return, investors will choose the less risky one. Investing is a tradeoff between risk and return; investors will take increased risk only if compensated by higher expected returns. Following this assumption Markowitz formulated the portfolio problem as that of finding the weighting of assets that minimise risk given a target expected return. Risk is measured as variance of expected returns.

The important message of MPT is that assets should not be selected only on characteristics that are unique to the asset. Rather, investors have to consider how each asset relates to all other assets.

To present the basic Markowitz mean-variance portfolio model, we need to introduce some notation. Let:

N be the number of assets available,
$\bar{r}_i$ be the expected (average, mean) return (per time period) of asset i,
ρ_{ij} be the correlation between the returns for assets i and j $(-1 \leq \rho_{ij} \leq +1)$,
s_i be the standard deviation in return for asset i,
σ_{ij} be the covariance between returns for assets i and j $(\sigma_{ij} = \rho_{ij} s_i s_j)$, and
M be the desired expected return from the portfolio chosen.

Then the decision variables are:

ω_i the proportion of the total investment associated with asset i $(0 \leq \omega_i \leq 1)$.

Observe that we imposed non-negativity (ω_i), meaning we can only *go long*, that is, buying and holding an asset in the hope that its price will rise. If we were to allow negative weights (so ω_i can be positive or negative), then we would be allowing shorting. Shorting (or short selling) is when investors borrow a particular asset and sell it immediately in the market, in the hope that the asset price will fall, enabling them to buy the asset back later at a lower price and return it to the original lender.

Using the standard Markowitz mean-variance approach we have that the portfolio optimisation problem is:

$$\text{minimise} \quad \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j \sigma_{ij} \tag{2.1}$$

subject to

$$\sum_{i=1}^{N} \omega_i \bar{r}_i = M, \tag{2.2}$$

$$\sum_{i=1}^{N} \omega_i = 1, \tag{2.3}$$

$$0 \leq \omega_i \leq 1, \quad i = 1, \ldots, N. \tag{2.4}$$

Here in Equation (2.1) we minimise the total variance (risk) associated with the portfolio. Equation (2.2) ensures that the portfolio has an expected return of M. Equation (2.3) ensures that the proportions sum to one, so that all available cash is invested in assets.

This formulation is a nonlinear programming problem. Usually nonlinear problems are difficult to solve, however in this case, since the objective is quadratic and $[\sigma_{ij}]$ is positive semidefinite (a property of covariance matrices), computationally effective algorithms exist so that in practice the above model can be solved with little difficulty.

The point of the above optimisation problem is to construct an *efficient frontier*, a smooth non-decreasing curve that gives the best possible tradeoff of risk against return, i.e. the curve represents the set of Pareto-*optimal* (*non-dominated*) portfolios.

Beasley (2013) gives one such efficient frontier, shown in Figure 2.1, for assets drawn from the UK Financial Times Stocks Exchange (FTSE) index of top 100 companies. Note how this nice smooth continuous curve runs from the minimum variance portfolio to the maximum return/maximum risk portfolio. Here we can choose to hold any of the portfolios on this efficient frontier. For this particular data set the minimum variance portfolio contained 30 out of the 100 assets.

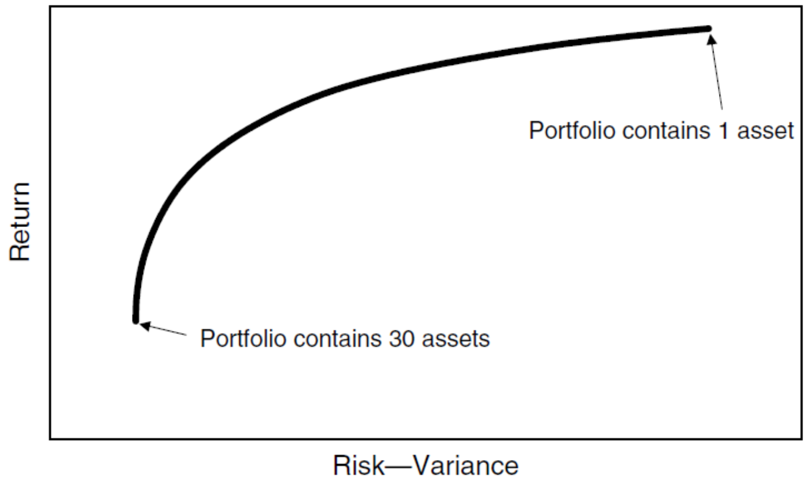

Figure 2.1: Efficient Frontier for the FTSE100

Based on Markowitz' mean-variance model, Treynor (1961), Sharpe (1964) and Lintner (1965) independently introduced the Capital Asset Pricing Model (CAPM). CAPM says that the expected return of an asset or portfolio equals the return on a risk-free asset plus a risk premium. If an asset is to be added to an already diversified portfolio, CAPM determines a theoretically appropriate rate of return that compensates the investor for taking the risk premium associated with that asset. CAPM assumes that each individual asset in a portfolio entails specific risk, but, through diversification, an investor's net exposure can be reduced to the systematic risk of the market portfolio.

The general idea behind CAPM is that investors need to be compensated in two ways: time value of money and risk. The time value of money is represented by the risk-free rate (which we denote as r_f), which means how much return an investor would expect from an absolutely risk-free investment over a given period of time. A rational investor that decides to take a risky investment expects at least to exceed the risk-free rate.

The other input to CAPM is the amount of compensation an investor needs for taking additional risk. This is calculated by taking the asset's sensitivity to non-diversifiable (specific) risk, often represented by the quantity β in the financial industry, and comparing the asset returns to the market premium (return over a risk-free investment). Given observed market returns r_m and asset a returns r_a, β_a is calculated as:

$$\beta_a = \frac{\sigma_{ma}}{\sigma_m^2}, \tag{2.5}$$

where σ_{ma} is the covariance between asset returns r_a and market returns r_m, and σ_m^2 is the variance of market returns r_m. Given expected market return $\bar{r}_m$, in order to decide whether an asset should be added to a portfolio we have to apply, according to CAPM, the formula:

$$\bar{r}_a = r_f + \beta_a(\bar{r}_m - r_f) \tag{2.6}$$

where asset a should be included in a portfolio only if its expected returns exceed $\bar{r}_a$.

An important assumption of CAPM is that asset prices move together because of one factor: the common movements of markets. The simplicity of CAPM led to the development of Arbitrage Pricing Theory (APT), first proposed by Ross (1976). APT considers that the expected return of an asset can be modelled as a linear function of various macroeconomic factors or theoretical market indices, where sensitivity to changes in each factor is represented by a factor-specific β coefficient. These models are usually referred to as factor models (see Wilmott (1998); Alexander (2001); Elton et al. (2007)). The standard form of a factor model can be written as

$$\bar{r}_a = \sum_{j=1}^{m} \beta_{aj} f_j + \epsilon_a \tag{2.7}$$

where β_{aj} are factor-specific sensitivities, m is the number of factors and ϵ_a is the portion of the return on asset a not related to any of the m factors.

The success of factor models in predicting returns depends on both the choice of the factors (f_j) and the method for estimating factor sensitivities (β_{aj}). Factors may be chosen according to economics (interest rates, inflation, etc.), finance (market indices, yield curves, exchange rates, etc.), fundamentals (book-to-market ratios, dividend yields, etc.) or statistics (factor analysis, principal component analysis, etc.). Sensitivities can be estimated using cross-sectional regression, time series techniques or eigenvalue methods.

The most popular factor model is the Fama-French three-factor model, designed by Fama & French (1993) to describe asset returns. Fama and French observed that two classes of assets have tended to outperform the market: small caps and assets with a high book-to-market ratio. They expanded CAPM to include portfolio exposure to these two classes. According to the Fama-French three-factor model, asset return is explained according to the formula:

$$\bar{r}_a = r_f + \beta_{a1}(\bar{r}_m - r_f) + \beta_{a2} \text{ SMB } + \beta_{a3} \text{ HML } + \epsilon_a \tag{2.8}$$

where SMB stands for "**S**mall (market capitalisation) **M**inus **B**ig" and HML stands for "**H**igh (book-to-market ratio) **M**inus **L**ow". Here β_{a1} is analogous to the classical β_a (given in Equation (2.5)) but not equal to it since there are now two additional factors.

2.2 Discussion on portfolio models

In the Markowitz framework, the portfolio is decided so as to minimise risk, where risk is defined as the in-sample variance in portfolio return. Clearly, risk can be defined in different ways. For example, a downside risk framework would equate risk with portfolio return falling below a predefined target. In this case, the objective of our optimisation model would be changed.

In fact, minimising risk is one of many possible objectives when defining a portfolio optimisation model. Take, for instance, the problem of designing an index tracking portfolio, where the objective is to replicate the performance of an index such as the S&P500 or the FTSE100.

In order to achieve this, full replication (buying all assets in the proportions that they compose the index) is possible, albeit for larger indices it can be an expensive strategy in terms of transaction cost. For example, whenever an asset enters/leave the index, the entire fund must be rebalanced, and any new money invested in the fund must be spread across all assets to mirror the index. For these reasons it is common not to adopt full replication. In such cases it is necessary to solve an index tracking portfolio optimisation model where the number of assets that can be bought is restricted. The objective of this problem is to minimise the tracking error, defined as the average squared difference between the tracking portfolio return and the index return.

Tracking error, however, is not the only way to measure the success of an index tracking portfolio. An alternative view on the above problem relates to regression. Suppose we perform a linear regression of the return from the tracking portfolio against the return of the index, i.e. the regression $r_t = \alpha + \beta R_t$, where r_t and R_t are the portfolio and index returns at time t. If we are looking for an index tracking portfolio then clearly we want an intercept $\alpha = 0$ and a slope $\beta = 1$. We can obtain α and β by using an ordinary least-squares regression (as in Canakgoz & Beasley (2009)), or, alternatively, by using quantile regression (Mezali & Beasley (2013)). The former looks for regression coefficients based on the least-squares mean regression line, while the latter uses coefficients based on median regression (the 50% quantile).

Recent work, other than that discussed above, dealing with index tracking can be found in Chen & Kwon (2012), Garcia et al. (2011), Guastaroba & Speranza (2012), Krink et al. (2009), Maringer (2008), Ruiz-Torrubiano & Suarez (2009), van Montfort et al. (2008), and Wang et al. (2012).

Essentially, different types of portfolios require different mathematical models, and even for portfolios intended for the same purpose, the model to use is not uniquely defined, such as the index tracking problem discussed above. In this thesis, we concentrate on three problems which in our view are not well defined in the literature: the problems of selecting an absolute return portfolio, a market neutral portfolio and an asset basket for an exchange-traded fund. We review the literature on each of these problems in the next sections.

2.3 Absolute return portfolios

The reader should be aware that the term 'Absolute Return Portfolio' is not clearly defined, as noted previously for example by Waring & Siegel (2006). Differing authors interpret the phrase 'Absolute Return' differently, as will be seen in our discussion of the literature below.

2.3.1 Stochastic programming models

One strand relevant to ARPs that can be found in the literature relates to guaranteed return funds. They fall within the ARP category as they aim to achieve a minimum absolute return. Work that deals with guaranteed return funds is often based on stochastic programming or some other form of future scenario prediction. The minimum return will hence be guaranteed provided the future is one of the predicted scenarios.

Dert & Oldenkamp (2000) proposed a stochastic programming model for a single-period guaranteed return portfolio that may include European put and call options. In this work a casino effect is shown to exist when one chooses portfolios to maximise expected return subject to achieving a minimum level of return under all circumstances (scenarios). The casino effect arises where there are high probabilities of obtaining low returns and low probabilities of receiving high returns. Since investors may dislike casino solutions the authors enhance their model by adding chance constraints which require that the probabilities of achieving returns less than pre-specified levels should be small. Numeric testing is based on options from the Standard & Poor's 500 index for 1997 with an investment horizon of 23 days. No details are given on the solution approach used.

Berkelaar et al. (2002) proposed an interior point approach based on primal-dual decomposition for a two-stage stochastic linear program. The work itself presents a general algorithm for this class of stochastic problems, but as an illustration their method is applied to a portfolio optimisation problem where an investor can invest in a money market account, a stock index and options on the index; moreover, a minimum return has to be guaranteed over a set of future scenarios. In their problem the portfolio (once constructed) can be rebalanced once (on a set date) before the end of the time horizon. The author mention, as advantages of their work, that the method proposed does not need a feasible starting solution and its computation time seems to grow linearly with the number of scenarios. Numeric testing is based on high liquidity options for the Standard & Poor's 500 index for 1999. The number of scenarios considered is 50 for the rebalancing date and 100 for the time horizon. No computation time for this portfolio optimisation problem is given, however there is a computational time comparison for some other test problems where their algorithm shows a much better performance than its deterministic equivalent. See Berkelaar et al. (2005) for an extension of this work to multistage stochastic convex programs.

Another work that relies on stochastic programming to guarantee a minimum return is that presented by Dempster et al. (2007). They proposed a stochastic formulation to a complex multivariable problem where, after an initial investment in a closed end guarantee fund, the objective is to hedge the risks involved in order to avoid having to buy costly insurance to guarantee the minimum return. This problem requires long-term forecasting in multiple time periods for many investment classes. They proposed a dynamic stochastic programming model to solve the problem. Stock prices are modelled using both standard geometric Brownian motion and geometric Brownian motion with Poisson jumps. Backtesting is presented for a 5 year period, from January 1999 to December 2003. The model is compared to the Euro Stoxx 50 index. Given a minimum barrier which the portfolio must exceed over time, the model behaves quite well, the only period where it drops below the barrier is on the 11th of September 2001. The number of scenarios considered is either 7776 or 8192, depending on the tree structure used for different horizon backtests, but no computation times are given. See also Dempster et al. (2006).

Herzog et al. (2007) applied sequential stochastic programming to an Asset Liability Management (ALM) problem that guarantees a minimal return on investments. The stochastic programming optimisation is resolved for every time interval on a new set of stochastic scenarios that is computed according to the latest conditional information. They show that such a technique approximates a continuous state dynamic programming algorithm and that, by using a sufficiently large number of scenarios, the difference between the exact solution and the approximation can be made arbitrarily small. They also argue that the most suitable risk measures for guaranteed return funds are shortfall risk measures. Hence, they define a penalty function relating to any shortfall below guaranteed return. Their objective function is to maximise multi-period return while keeping the penalty function under a certain level as defined by a coefficient of risk aversion. The model also includes transaction costs. They presented a case study relating to a Swiss fund with quarterly data over the period 1988 to 2005 with up to 5000 scenarios.

Barro & Canestrelli (2010) proposed a multistage stochastic programming framework for a dynamic asset allocation problem which takes into account the conflicting objectives of a minimum guaranteed return and of an upside capture of asset returns. They argue that maximising the upside capture increases the total risk of the portfolio, thus they attempt to balance this by introducing a second goal where they try to minimise the shortfall with respect to the minimum guarantee level. To combine these two conflicting goals they formulate them in the framework of a double dynamic tracking error problem using asymmetric tracking measures, one for a risky benchmark and one for a minimum guaranteed benchmark. The objective is a combination of both tracking errors functions. They describe the uncertainty of future returns by using the concept of a scenario tree where each scenario is represented as a path from the origin to a leaf of the tree. An interesting feature of this model is the introduction of liquidity constraints which take into account the bid-ask spread. They also briefly discuss a second approach where the problem of minimum guaranteed return is tackled with the introduction of chance constraints. No numerical results are given.

2.3.2 Other models

There are also papers presented in the literature that (unlike those discussed above) do not use stochastic programming.

Nishiyama (2001) considered an absolute return strategy derived from multi-manager investment, a fund of funds (FoF) approach, in Japan. He argued that FoFs have traditionally low correlation against the benchmark index and little impact from external changes, thus being absolute return strategies. He observes that nonlinear events occur frequently in the market, e.g. crashes, and that this phenomena should not be interpreted using traditional theory framework, which divides a portfolio risk in two: systematic (market) and unsystematic (specific) risk. He then proposes the use of the physics theory of Self-Organised Criticality (SOC) to find the point at which a system changes its behaviour or structure, for instance, from solid ice to liquid water. Unlike the melting case, where the control parameter is the temperature, a SOC reaches a critical state by its intrinsic dynamics, e.g. the market reaches a critical state when it crashes and this is not due to a control parameter. To understand the dynamics he focused on the correlation matrix of nonlinear fluctuations in the market, where he looked into historical movement of asset correlation and assumed its behaviour would be ex-ante signals of a market crisis. He introduces a third risk category called group risk, which lies between systematic and unsystematic risk and accounts for the nonlinear type of market movements. He believes that minimising variance is not sufficient to measure risk, and he emphasises the importance of correlation rather than variance. Simulated results over the period 1995-2000, so including the 1998 Russian crisis and the failure of Long-Term Capital Management, were presented.

Korn (2005) proposed a different approach for portfolio selection with a positive lower bound on the final wealth. The solution given consists of transforming the original problem into an equivalent unconstrained portfolio problem with a modified utility function that does not include the lower bound. Unlike the majority of work on guaranteed return funds, time is considered a continuous series instead of being divided into discrete intervals. As mentioned by the author, the unconstrained version is solved analytically via numerical methods, although no details are given on which numerical methods were used. After the problem is solved, the optimal final wealth is separated into a hedging term, needed to satisfy the requirement of a minimum final wealth, and a speculative term. The hedging term is a portfolio made up of put options and stocks. The speculative part could be calculated by computing the delta of the corresponding options. A few examples are given that demonstrate the relationship between stock investment and the growth or decay of total wealth, however, the deltas (speculative term) are not computed since they require lengthy expressions to be solved. No computational results are given for real world data. Stock prices are modelled using generalised geometric Brownian motion.

Amenc et al. (2008) proposed an approach based on a dynamic core-satellite portfolio. This technique, explained in detail in Amenc et al. (2004), consists of splitting the cash allocation into different portfolios. The core portfolio is mainly a low-risk portfolio that intends to respect the investor's long-term risk return objectives, while the satellite portfolio provides access to upside potential by investing in more risky assets that are expected to outperform the benchmark. At discrete time intervals, the investor decides the proportion invested in each portfolio based on a minimum guaranteed value that is relative to the benchmark, e.g. 90% of an underlying index. The dynamic allocation process will systematically increase the exposure to the satellite portfolio when it does well with respect to the core, while controlling risks by shifting to the core when the satellite does poorly. They gave an example where the core is composed of Euro bonds and the satellite is an exchange-traded fund relating to the Euro Stoxx 50. They compared their core-satellite approach with an active manager simulation, and showed that overall their dynamic core-satellite portfolio led to better results than a simulation of an actively managed portfolio built on largely accurate forecasts.

Lejeune (2011) considered an absolute return strategy derived from a long-only fund of funds approach which was formulated as a mixed-integer nonlinear programming problem. Portfolio variance is constrained to be below a given limit, this is ensured through a Value-at-Risk (VaR) constraint that limits the magnitude of the loss with a specified probability level over a certain period of time. In the model presented VaR takes the form of a probabilistic constraint. They estimate asset returns by using the Black-Litterman framework (Black & Litterman (1992)) which attempts to overcome problems of highly-concentrated input-sensitive portfolios. The model presented is computationally expensive and thus the probability constraint is approximated deterministically by a second-order cone constraint which makes the problem convex for a wide range of probability distributions. They present a specialised nonlinear branch-and-bound algorithm which is implemented by an open source nonlinear solver. The branch-and-bound is compared to two other solvers in terms of computational performance over 12 different problem instances.

Zymler et al. (2011) proposed an approach based on combining robust optimisation with options, an approach they call insured robust portfolio optimisation. Robust optimisation (e.g. see Ben-Tal & Nemirovski (1998)) gives a guarantee provided data variation lies within a specified uncertainty set. They add another layer of guarantee to hedge against rare events which are not captured by a reasonably sized uncertainty set. They enrich the portfolio with specific derivative products to obtain a deterministic lower bound that essentially provides a barrier (insurance) such that the portfolio value cannot drop below the required level. They argue that enlarging the uncertainty set to cover extreme events (instead of adding the second layer of insurance) would lead to robust portfolios that are too conservative and could perform poorly under normal market conditions. The model they develop is a convex second-order cone program which is scalable in the number of stocks. Numeric results were given based on simulated data as well as historical data, where they observed that while the uninsured portfolio tends to have higher expected returns in normal market conditions, the proposed insured model shows clear advantages in terms of Sharpe ratios, expected returns and cumulative wealth when the markets behave abnormally.

2.4 Market neutral portfolios

The history of Market Neutral Portfolios traces back to the mid-1980s, when Morgan Stanley assembled a group of mathematicians, physicists and computer scientists to develop trading programs whose intention was to take the intuition and trader's "skill" out of arbitrage and replace it with disciplined, consistent filter rules. Among other things, this group identified pairs of securities that tended to move together. This marked the beginning of *Pairs trading*, which is the first known market neutral technique and it is generally assumed to be the ancestor of statistical arbitrage (see Gatev et al. (2006)).

Pairs trading works as follows. If two assets (say P and Q) are in the same industry or have similar characteristics, one expects the two asset returns to track each other within a certain error. If P_t and Q_t denote the corresponding price series which are historically correlated, then we can model a system such as:

$$\ln(P_t/P_{t_0}) = S\ln(Q_t/Q_{t_0}) + \epsilon_t \tag{2.9}$$

where ϵ_t is a stationary, or mean reverting, process, usually known as *cointegration residual* and S is a constant equating $\ln(P_t/P_{t_0})$ with $\ln(Q_t/Q_{t_0})$. The model above suggests an investment strategy where, if ϵ_t is sufficiently positive, we short $S£$ of asset Q for every $1£$ invested in long positions of asset P. Conversely, if ϵ_t is sufficiently negative, we go short on P and long on Q. The portfolio is expected to oscillate near some statistical equilibrium, bringing ϵ_t closer to zero. This is usually called a mean-reversion strategy as the investor bets that the prices will eventually revert to their historical trends. This is typically associated with market overreaction: assets are temporarily under or overpriced with respect to one or several reference securities (Lo & MacKinlay (1990)). See Pole (2008) for a comprehensive review on statistical arbitrage and cointegration.

Pairs trading is considered market neutral as it provides a hedge against market risk. For example, if the whole market crashes, and the two assets plummet along with it, the trade should result in a gain on the short position and a loss on the long position, leaving the profit close to zero in spite of the large move. Note, however, that market neutral investing is not a single strategy; pairs trading is one possible technique to achieve market neutrality. For instance, an alternative strategy is known as delta neutral, where delta is defined as the sensitivity of an option value with respect to changes in the underlying asset's price when all other variables remain unchanged. A delta neutral portfolio is one which tries to maintain its value unchanged when small changes occur in the value of the underlying securities. Such a portfolio typically contains options and their corresponding underlying securities such that positive delta components (namely, long call or short put options) and negative delta components (short call or long put options) offset. Work found in literature that study market neutral investments are summarised below.

Avellaneda & Lee (2010) define a market neutral portfolio as one that is uncorrelated with the market. They define the market as the combination of multiple factors, where they present a model to explain stock returns as composed of a systematic component dependent on factors and an idiosyncratic (uncorrelated) component. They then regard a market neutral portfolio as one where the portfolio has zero exposure to these factors. To estimate the factors they use two different approaches. The first one is called the Principal Components Analysis (PCA) (see Jolliffe (1986)) which uses historical data to create an empirical correlation matrix. From this matrix they extract eigenvalues and rank them in decreasing order, where those that exceed a given percentage of the trace of the correlation matrix are considered significant. From the significant eigenvalues they form weighted "eigenportfolios", which are the estimates of factors. In the second approach, they select a sufficiently diverse set of Exchange-Traded Funds (ETFs) and consider them as factors. Extensive computational results were given, where they noted that the performance of mean-reversion strategies appears to benefit from situations where most of the variance can be "explained" (with significant regression coefficients) by a relatively small number of factors. If the "true" number of factors needed to explain variance was very large then using only a few factors would not be enough to "defactor the returns", so residuals would "contain" market information that the model is not able to detect. If, on the other hand, they used a large number of factors, the corresponding residuals would have small variance, and thus the opportunity of making money, especially in the presence of transaction costs, is diminished.

Baronyan et al. (2010) investigated different pairs trading strategies by combining 7 different policies of pairs selection with 2 trading methods, in a total of 14 strategies. Together the pairs are meant to be "market neutral" although, as is clear from the varying strategies proposed in Baronyan et al. (2010), there are many different measures that can be used to decide whether a pair is "market neutral" or not. Computational results were presented for pairs of stocks drawn from the Dow Jones 30 index, so $\binom{30}{2} = 435$ pairs. Yearly tests were performed from $N = 1999, \ldots, 2006$, where years N and $N+1$ comprise the training (in-sample) period, and year $N+2$ comprises the testing (out-of-sample) period. Thus, the last test included the year of the global financial crisis. They compared the seven different strategies and found that in general all of them resulted in positive cumulative returns, especially in 2008. They argued that mispricings in pairs of similar stocks are more commonplace in a global crisis, thus allowing more trading possibilities to emerge in bad times.

Ganesan (2011) used a regression of individual stock returns against a number of factors. He defines a market neutral portfolio as one whose exposure to factors is zero. He presents a single factor and a multi-factor model for describing the stocks expected returns. He adopted a Markowitz mean-variance approach to portfolio construction, where new constraints were added to transform the problem into that of finding a market neutral portfolio. Using geometrical subspace analysis he shows that any portfolio can be decomposed as the sum of a market-neutral portfolio and a factor-exposure portfolio. Examples with three stocks were used to explain both the single factor and multi-factor approaches. An empirical study was performed comprising all U.S. stocks traded on New York and NASDAQ (about 2000 stocks) during the period from 1998 to 2010. The study results show that the performance of the market-neutral component depends on the cross-sectional variation of stock returns (called dispersion), while the performance of the factor component depends on the volatility of the overall stock market.

Kwan (1999) used a regression of stock returns against market returns and regarded a market neutral portfolio as one where the long portfolio weighted regression parameters relating to slope are equal to the equivalent short parameters. His model focused on accurately portraying institutional procedures for short selling (i.e. by including in the model all possible costs involved in short transactions) while adhering to the mean-variance framework. The portfolio is subject to a constraint where the weighted average of the long and short beta (slope) coefficients must be equal and the model's objective is to maximise the Sharpe ratio (Sharpe (1966, 1975, 1994)). The securities are doubled (to account for long and short) and ranked according to criteria that take into consideration how undervalued (overvalued) a particular asset is. The model is solved via an iterative procedure where securities are added one by one to the portfolio while maintaining feasibility. Computational results were given for one illustrative example involving 20 stocks from the Dow Jones Industrial index. The model is flexible enough to accommodate different market outlooks by adjusting the weights between the market-neutral and market-sensitive components.

Ma et al. (2011) used a regression of individual stock returns against a number of factors. In their regression different parameters applied depending upon the market regime (e.g. bull or bear market, where prices are increasing or decreasing respectively). They formulated a stochastic linear program to maximise portfolio return whilst constraining the portfolio factor exposure to lie within limits. The market neutral strategy is implemented by constructing and rebalancing the portfolio that has overall zero betas for all relevant risk factors and thus the return of the portfolio under such a strategy is uncorrelated with the market risk factors. They used Bayesian information criteria to estimate the number of regimes. Three regimes were identified, where the third (apart from bull and bear) was a transitional market. A transition probability matrix between the three regimes is given, but no details are provided on how it was estimated. Computational results were given for one example involving nine sector ETFs over the period January 2005 to September 2009. The strategy is compared to a benchmark strategy which invests passively and equally among the nine sector ETFs. Their results show that, in general, the regime-dependent strategy outperforms the benchmark strategy. Their evaluation, however, did not include a bear market situation.

Pai & Michel (2012) used a regression of stock return against market return and regarded a market neutral portfolio as one where the portfolio weighted regression parameters relating to slope are nonlinearly constrained. They used a Markowitz mean-variance approach to portfolio construction with a nonlinear portfolio risk constraint. They argued that while the problem of constructing a basic market neutral portfolio could be modelled as a linear program, its complexity would be greatly increased by adding constraints such as investor preferences, market norms or investment strategies. Thus, they implement a differential evolution heuristic (Storn & Price (1997)) that exploits a penalty function strategy and employs weight standardisation procedures. These procedures are responsible for the complex constraints handling, ensuring the feasibility of the population of individuals in each step of the evolution cycle and leading to faster convergence. Computational results for portfolios with up to 64 stocks drawn from the Bombay stock exchange were given, where they performed statistical hypothesis tests to prove the robustness of their out-of-sample results.

Badrinath & Gubellini (2011) examined 27 market neutral funds over the time period October 1990 to December 2007. They consider a market neutral strategy as a specific implementation of a long-short strategy that minimises exposures along one of multiple possible dimensions: VAR-neutral, mean-variance neutral, dollar-neutral just to name a few. They concluded that market neutral funds monthly returns were uncorrelated with those of the market, where the market is represented by the Fama-French three-factor model and its momentum augmented version, the Carhart four-factor model (see Carhart (1997)). They also conducted an evaluation of portfolio performance, where they concluded that market neutral funds require relatively frequent adjustments to market-risk exposure to achieve their goals; their analysis also showed a superior risk-adjusted performance in down-market states when compared to up-market states.

During a week in August 2007, a number of high-profile market neutral hedge funds experienced unprecedented losses as the credit crunch crisis hit financial markets. Khandani & Lo (2007, 2011) discussed the effect of these events on long/short market neutral funds. They also attempted to explain the causes that led to such unusual market movements. In Khandani & Lo (2007), they concluded that the rapid unwinding (liquidation) of such a fund may have led to a cascade effect. They argue that these events are not particularly relevant to the general efficacy of quantitative investing since the losses were more likely to be the result of a firesale rather than shortcomings of quantitative methods. However, they explain that the 2007 events show that problems in one corner of the financial market can spill over to a completely unrelated corner, leading them to discuss regulation of the hedge-fund sector. In Khandani & Lo (2011) they identified indirect evidence of two specific unwinds on August 1st and 6th 2007. They simulated the performance of a high-frequency mean-reversion strategy that indirectly explained why liquidity declined sharply during August 2007.

Patton (2009) pointed to a lack of clarity in the meaning of the term "market neutral" and considered a number of different definitions. He considered the concept of neutrality more generally than that implied by the use of beta by proposing five alternative concepts: mean-neutrality (which nests the correlation or beta-based definition), variance, Value-at-Risk and tail neutrality, and finally a concept of complete neutrality which corresponds to statistical independence of fund and market returns. Statistical tests for each concept of neutrality were introduced in the hope of aiding investors' evaluation of funds. A detailed study of a combined database of 1423 hedge funds in a variety of fund styles was performed, using monthly returns, over the period April 1993 to April 2003. The market benchmark was considered to be the S&P 500 index for most of the hedge funds, and he showed that the results do not vary greatly if other equity indices are used. He found that approximately 28% of 197 funds described as market neutral exhibited significant correlation with the market at the 5% significance level. When comparing to other fund categories he argues that his findings suggest that many market neutral funds are in fact *not* market neutral, but overall, at least, they are more market neutral than other categories.

2.5 Exchange-traded funds

Exchange-Traded Funds (ETFs) were introduced in the 1990s, early issues around their introduction are discussed in Kupiec (1990) and Gastineau (2001). Kupiec discussed ETFs predecessors, which were called Index Participation shares (IPs) and had been recently approved but not yet issued. Their goal was to provide investors with a flexible tool to trade an entire portfolio in a single transaction. The idea of a portfolio traded as a share actually dates back from the 1970s and 1980s, when the introduction of S&P 500 index futures provided an arbitrage link between futures contracts and the traded portfolio of stocks. The effect of all these developments was to make portfolio trading in either cash or futures markets an attractive activity for many trading desks and for many institutional investors, which naturally led to an interest in a readily tradable portfolio or basket product for smaller institutions and individual investors. In the early 1990s IPs grew in popularity. IPs were much like a futures contract, but they were margined and collateralised like stocks. Like futures, there was a short position for every long and a long position for every short. A federal court in Chicago found that the IPs were indeed illegal futures contracts and had to be traded on a futures exchange. This eventually led to the end of IPs and paved the way for the development of ETFs and their subsequent introduction in 1993.

Poterba & Shoven (2002) provided some statistics on the growth of ETFs since their introduction. The total ETF market was approximately US$80bn in 2001, having grown steadily since 1993. Two ETFs (the SPDR Trust SPY and the Nasdaq-100 QQQ) made up some 60% of the market at that time. They compared the pre-tax returns of the SPDR Trust SPY and the Vanguard Index 500 Fund, which is a domestic equity index fund, to the S&P 500 returns. The calculations showed that the average return on both funds was close to returns from the S&P 500 index, with the Vanguard Index 500 Fund returns being slightly higher. They noted that the fact that ETF shares values are detached from the actual basket value can lead to non-trivial year to year return differences. They also showed that the difference in returns between the two funds was reduced when after tax returns were considered, mainly because of the higher tax burden associated with mutual and index funds.

Boehmer & Boehmer (2003) considered the introduction by the New York Stock Exchange (NYSE) of trading in three large ETFs (SPY, QQQ and a Dow Jones ETF, DIA), plus a number of smaller ETFs, that had previously been traded just on other exchanges. They documented double-digit percentage declines in quoted, effective, and realised spreads after the NYSE entry. The difference between effective and realised spread, an aspect of liquidity, also decreased significantly. The NYSE entry considerably improved liquidity in the entire market and also in the individual market centres. Detailed tests were conducted showing that this result was not due to shifts in informed trading or a temporary phenomenon. They also concluded ETF trading costs were lowered. A possible explanation for this reduction rests on the assumption that different market centres have comparative advantages with certain order types. For example, one market may be better able to handle a large volume of small, uninformed orders, while another may be better able to handle a high volume of large orders, because it has a deeper pool of liquidity. Under this view, the NYSE entry may have led to a more efficient allocation of orders to the respective lowest-cost market centre, such that all markets are able to offer lower trading costs.

Kostovetsky (2003) examined the conditions under which it is preferable for an investor to invest in an (index tracking) ETF as compared with a conventional index tracking mutual fund. He developed a one-period model, which is then expanded to multi-period, to examine the major differences between ETFs and index funds. The model, albeit simple, emphasised the importance of management fees, shareholder transaction fees and taxation. He also discusses qualitative differences between ETFs and index funds. For example, some advantages of ETFs are the convenience of being able to trade at any time of the day and the ability to buy on margin and to short sell. The latter makes ETFs appropriate for hedging strategies.

Alexander & Barbosa (2008) examined the hedging problem which arises in ETF creation and redemption when the portfolio underlying the ETF shares involves illiquid stocks with relatively high transaction costs. Their work examined the use of minimum variance hedging using three different performance criteria: aversion to negative skewness, excess kurtosis and effective reduction in variance. They found that minimum variance hedging was preferable to a simple hedge (based on one long position in the ETF and one short position in futures) if aversion to negative skewness and positive excess kurtosis were considered. Their results considered an out-of-sample period from January 2001 until September 2006, in which they identified three distinct regimes. They argue that the performance of each hedging strategy was independent of the market regime as little difference in performance was observed in different regimes.

Mariani et al. (2009) examined the return distributions of three ETFs and their corresponding benchmark indices using a Levy model. They described the temporal evolution of financial markets as a normalised Truncated Levy Flight (TLF), which in their view is more suitable for long-range correlation scales than classical Levy models. They examined the S&P 500 SPDR, the Dow Diamonds and the PowerShares QQQ and compared them with the behaviour of their indices, namely the S&P 500, the Dow Jones Industrial Average and the NASDAQ 100, respectively. The time period considered is extensive as their data is composed of daily ETF and index prices from when the respective ETF was issued until October 2007. The S&P 500 SPDR, for example, was issued in 1993. They concluded that these ETFs exhibited the same behaviour as their indices and argued that the normalised TLF model allowed them to accurately complete a numerical analysis.

Avellaneda & Zhang (2010), Giese (2010), Haugh (2011) and Jarrow (2010) all considered ETFs from the perspective that the underlying price dynamics of the assets can be modelled using some stochastic process (e.g. Brownian motion).

Avellaneda & Zhang (2010) presented an exact formula linking the return of a leveraged ETF (an ETF which hopes to achieve a multiple of the benchmark return, for example, a $2\times$ leveraged ETF attempts to achieve twice the daily return of its benchmark) with the corresponding multiple of the return of the unleveraged ETF and its realised variance. They tested their formula using a number of ETFs (twenty-two $2\times$ ETFs, twenty-two $-2\times$ ETFs, six $3\times$ ETFs, six $-3\times$ ETFs) and concluded that their formula is a good explanation of ETF price behaviour. Their study showed that leveraged funds could be used to replicate the returns of the underlying index, provided a dynamic rebalancing strategy was used. Empirically, they found that rebalancing frequencies required to achieve this goal are on the order of one week between rebalancings. From their formula they draw a series of conclusions about leveraged ETFs. For example, if the price of an underlying unleveraged ETF does not change significantly over time, but the realised volatility is large, the leveraged ETF will underperform the corresponding multiplied return of the unleveraged ETF. This make leveraged ETFs unsuitable for buy-and-hold investors.

Giese (2010) presented a model in which there is a tradeoff between exploiting the potential of higher returns, which grow linearly with the ETF leverage factor (e.g. 2×, 3×), and adverse losses owing to the volatility of the underlying, which is proportional to the ETF leverage squared. Their model seeks an optimal leverage value, and it can be adapted to either long or short leveraged trading strategies, but not both at the same time. Leveraged ETFs are rebalanced on a daily basis and hence transaction costs cannot be neglected, so these are taken into account. He observed that the optimal leverage value strongly depends on prevailing market conditions, such as a bullish or bearish market. They considered a numeric example based on the EUROSTOXX 50 total return index in two different time periods, from 1991 until 2007 (when at the end the markets were close to a peak) and from 1991 to 2009 (when at the end the markets were in a recession). Their optimal leverage model outperformed both a 2× ETF and 4× ETF simulation.

Haugh (2011) considered a constant proportion trading strategy, where the fraction of the total wealth invested in a risky asset remains fixed and does not vary over time. Such a strategy requires constant (daily) rebalancing. He argues that this strategy can be used to explain the performance of leveraged ETFs. They presented the terminal wealth of a constant proportion trading strategy as a function of terminal asset prices, which they used to explain leveraged ETF performance when specialised to the case of just one underlying asset. Hence, a leveraged ETF that tracks an index (composed of multiple assets) was not considered. They argued that an actively managed constant proportion ETF could be a suitable product for investors, although the costs associated with daily rebalancing would be prohibitively expensive for any small and individual investors. In addition, the manager of a constant proportion ETF (or a normal leveraged ETF) would necessarily sell at the close after an up-day and buy at the close after a down-day and would therefore tend to dampen market volatility. Because the direction of the daily rebalancing trades are widely known in the market, it is suspected that many proprietary trading desks illegally take advantage of advance knowledge of pending orders to profit from these trades. He suggests less frequent rebalances as a way to avoid this risk and incidentally reduce management costs at the expense of rendering a less useful approximation.

Jarrow (2010) presented a model where investment is dynamically switched between an ETF (whose value followed a diffusion process) and a money market account in an attempt to achieve a given multiple of ETF return. The model enables one to characterise the return distribution of the leveraged ETF over any investment horizon. The instantaneous return on a k-times leveraged ETF is equal to k times the return on the ETF less the interest paid on the borrowings. He also shows that the k-times return does not hold over any finite investment horizon, due to the interest rate reduction component. He emphasises that promotional materials for leveraged ETFs warn about the volatility effect that devalues leveraged ETFs but usually ignore interest rate reductions. No numeric results were given for their model.

2.6 Conclusion

In this chapter we gave an extensive review of previous studies on portfolio optimisation models. In Section 2.1, we presented a historical context for portfolio optimisation. In Section 2.2 we mentioned that portfolios are very diverse, with many different objectives. Even for portfolios intended for the same purpose, the model to use is not uniquely defined, which makes it difficult to compare different works in the literature.

In Section 2.3 we discussed several works that define their strategies as absolute return portfolios. Most make use of a minimum guaranteed return and are based on stochastic programming, but other solution methods also exist. In Section 2.4, we presented several works that deal with market neutral portfolios. Market neutral models are more clearly defined than absolute return portfolios, albeit there are still multiple ways to define what market neutrality is.

Finally, in Section 2.5, we discussed a brief history of exchange-traded funds and related works. Optimisation models for ETFs are still not common in the literature, and most works study ETF properties and performance.

Overall, we summarise the works on these three models as very fragmented, with different models and different data result in isolated papers, with great difficulty in connecting them in a mathematical/data sense. Many works do not give detailed computational results.

Chapter 3

Absolute return portfolios

3.1 Introduction

Absolute Return Portfolios (henceforth ARPs) are financial portfolios that aim to produce a good return regardless of how the underlying market performs. This (clearly) is a relatively easy task when the market is performing well, a much less easy task when the market is performing poorly. Essentially investors are interested in ARPs either because:

- they believe that the market will perform poorly, and so wish to focus on portfolios that will not perform as poorly; or
- they are unsure of how the market will perform and wish to hold an ARP as insurance against market deterioration.

ARPs are a relatively popular strategy amongst managers of some hedge funds, which, as their name suggests, often seek to hedge some of the risks inherent in their investments using a variety of methods. Their objective is to achieve absolute returns by balancing investment opportunities with the risk of financial loss. Al-Sharkas (2005), Connor & Lasarte (2010), Jawadi & Khanniche (2012) and Till & Eagleeye (2003), discuss the various strategies that hedge funds can adopt.

ARPs are sometimes called market neutral portfolios as they are designed to have a low correlation with overall market return. Whilst, due to this strategy, ARPs may be able to achieve positive returns in falling markets, on the other hand they may not perform as well as market indices or other types of investments in rising markets. However, the fear of significant financial events (we have seen the 2008 subprime financial crisis; in the near future will we see a Eurozone default?) makes ARPs popular amongst investors, who see them as a reasonable strategy to adopt given market uncertainty and volatility.

In this chapter, we present a three-stage mixed-integer zero-one program for the problem of designing an ARP. Our formulation includes transaction costs associated with trading, a constraint limiting the number of assets that can be held and a limit on the total transaction costs that can be incurred. The first two stages relate to a regression of portfolio return against time, whilst the third stage relates to minimising transaction cost. ***One feature of note in our ARP approach is that we do not specify the return that the ARP should achieve; rather that emerges as a result of an optimisation.***

The original contribution of our model/formulation relates not to the constraints adopted (which are in fact standard and have been seen before in the literature, e.g. in Canakgoz & Beasley (2009)). Rather the original contribution of our model relates to a clear definition of an ARP via the three-stage objective function.

Because our approach is flexible we are able to extend it to the problem of designing portfolios with differing characteristics. In particular we present models for enhanced indexation (relative return) portfolios and for portfolios that are a mix of absolute and relative return.

The rest of this chapter is organised as follows. In Section 3.2 we present our regression based three-stage mixed-integer zero-one program used to decide an ARP. In Section 3.3 we go on to show how this formulation can be extended to design portfolios with differing characteristics.In Section 3.4 we present computational results for portfolios derived from universes defined by S&P international equity indices. In Section 3.5 we present our conclusions.

3.2 Problem formulation

3.2.1 Overview

In the formulation presented in this chapter we adopt the view that in seeking an ARP we are seeking a portfolio that ***achieves a constant return per time period***. Of course we may not find a portfolio with this property - but in terms of what we desire our view is that if we can find such a portfolio then we would have an ideal ARP - giving us the same (constant) return in each and every time period. Here the notion that an ARP is somehow 'disconnected' from the market is captured by the constancy of return. This is because if we achieve a constant return in each and every time period, when (presumably) the market varies, how can the portfolio and the market be related? This obviously simplifies the situation, but does reflect the essence of what we would like to achieve in an ARP, a portfolio with a constant return per time period.

Now if we desire a portfolio with a constant return per time period should we specify what that constant return is, or should we allow it to be determined in some other fashion? Specifying the desired level of constant return might at first sight seem attractive, but in reality it has some difficulties. If we specify a value that is too low then we may choose a portfolio that will not generate as much return as could otherwise be achieved. If we specify a value that is too high then we may not be able to find a portfolio that achieves that return (even in-sample). Because of these considerations we in our model ***do not specify the return that the ARP should achieve; rather that emerges as a result of an optimisation***.

In this chapter we adopt a regression based view of the problem of selecting an ARP. A key computational advantage of this approach is that it allows us to develop a nonlinear formulation which can be linearised in a standard way. Our approach is a three-stage mixed-integer zero-one program. As such standard software packages, such as CPLEX Optimizer (2013), can be used to find optimal solutions. Computational experience reported in this chapter is that, for the test problems we examined, optimal solutions can be found very quickly.

Before presenting our model/formulation we should mention here the well-known regression based models, discussed in Section 2.1, that relate asset return (and by implication/extension portfolio return) to various factors, for example the capital asset pricing model (Sharpe (1964)), the Fama-French three factor model (Fama & French (1993, 1996)) and the Carhart four factor model (Carhart (1997)). Recall here that, as discussed above, we are ***defining*** an ARP as a portfolio that (ideally) achieves a constant return per time period. As such a regression of portfolio return against time is the appropriate regression to use. ***Regressing portfolio return against other factors (as in these models) would not satisfy the definition we have set out for an ARP.***

In the following sections we give our notation and present the constraints and objective that we use to find an ARP.

3.2.2 Notation

We observe over times $t = 0, 1, 2, \ldots, T$ the value of N assets. We are interested in selecting, at time T, the best set of K assets to hold (where $K < N$), as well as their appropriate quantities (number of asset shares or, equivalently, units). Let:

T	be the time period where the composition of the portfolio is decided
V_{it}	be the value (price) of one unit of asset i at time t

r_{it}	be the single period continuous time return for asset i at time t, i.e. $r_{it} = \ln(V_{it}/V_{it-1})$
$\hat{\alpha}_i$,$\hat{\beta}_i$	be the least-squares regression intercept and slope for asset i when the returns r_{it} from asset i are regressed against time, i.e. the regression equation is $r_{it} = \hat{\alpha}_i + \hat{\beta}_i t$
I_t	be the value of the index at time t
R_t	be the single period continuous time return for the index at time t, i.e. $R_t = \ln(I_t/I_{t-1})$
$\hat{A}$,$\hat{B}$	be the least-squares regression intercept and slope when the returns R_t from the index are regressed against time, i.e. the regression equation is $R_t = \hat{A} + \hat{B}t$
C	be the total value (≥ 0) of the current ARP $[X_i]$ at time T, $\sum_{i=1}^{n} X_i V_{iT}$, plus cash change (either new cash to be invested or cash to be taken out)
f_i^b	be the fractional cost of buying one unit of asset i at time T, so that the cost incurred in buying one unit of asset i at time T is $f_i^b V_{iT}$
f_i^s	be the fractional cost of selling one unit of asset i at time T, so that the cost incurred in selling one unit of asset i at time T is $f_i^s V_{iT}$
γ	be the limit ($0 \leq \gamma \leq 1$) on the proportion of C that can be consumed by transaction cost
ε_i	be the minimum proportion of the ARP that must be held in asset i if any of the asset is held
δ_i	be the maximum proportion of the ARP that can be held in asset i

Then our decision variables are:

x_i	the number of units (≥ 0) of asset i that we choose to hold in the ARP
w_i	the proportion of the initial investment (cash) C held in asset i ($0 \leq w_i \leq 1$)
z_i	$\begin{cases} 1 & \text{if any of asset } i \text{ is held in the ARP} \\ 0 & \text{otherwise} \end{cases}$
G_i	the transaction cost (≥ 0) associated with buying or selling asset i

Without significant loss of generality (since the sums of money involved are large) we allow $[x_i]$ to take fractional values. Note here that as $x_i \geq 0$ we are not considering short-selling, rather we are considering long-only portfolios. Note also that the transaction costs and the proportion variables are strictly unnecessary since they can be substituted out algebraically and represented as functions of the x variables, but are introduced here to ease the mathematics presented.

3.2.3 Constraints

The constraints associated with the ARP problem are:

$$w_i = x_i V_{iT}/C, \quad i = 1, \ldots, N \tag{3.1}$$

$$\varepsilon_i z_i \leq w_i \leq \delta_i z_i, \quad i = 1, \ldots, N \tag{3.2}$$

$$\sum_{i=1}^{N} z_i = K \tag{3.3}$$

$$G_i \geq f_i^s (X_i - x_i) V_{iT}, \quad i = 1, \ldots, N \tag{3.4}$$

$$G_i \geq f_i^b (x_i - X_i) V_{iT}, \quad i = 1, \ldots, N \tag{3.5}$$

$$\sum_{i=1}^{N} G_i \leq \gamma C \tag{3.6}$$

$$\sum_{i=1}^{N} x_i V_{iT} = C - \sum_{i=1}^{N} G_i \tag{3.7}$$

$$G_i \geq 0, \quad i = 1, \ldots, N \tag{3.8}$$

$$x_i \geq 0, \quad i = 1, \ldots, N \tag{3.9}$$

$$z_i \in [0, 1], \quad i = 1, \ldots, N \tag{3.10}$$

Equation (3.1) defines the proportion w_i of the ARP invested in asset i. Equation (3.2) ensures that if an asset i is not in the ARP ($z_i = 0$) then w_i is zero (and Equation (3.1) then ensures that x_i is also zero). Equation (3.2) also ensures that if an asset i is chosen to be in the ARP ($z_i = 1$) the amount of the asset held satisfies the proportion limits defined. Equation (3.3) ensures that there are exactly K assets in the ARP. Equation (3.4) defines the transaction cost associated with selling asset i, where we have sold the asset if the current holding X_i is greater than the new holding x_i. Equation (3.5) defines the transaction cost associated with buying asset i, where we have bought the asset if the new holding x_i is greater than the current holding X_i. Equation (3.6) limits the total transaction costs. Equation (3.7) is a balance constraint which ensures that the value of the ARP after trading is equal to its value before trading minus the total transaction costs incurred. Equations (3.8) and (3.9) are the non-negativity constraints whilst Equation (3.10) is the integrality constraint.

Observe that since variables $[x_i]$ can take fractional values the value of the portfolio after rebalancing is exactly equal to C minus transaction costs, i.e. there is no leftover cash after the portfolio is rebalanced.

3.2.4 Three-stage objective

In this section we give our three-stage objective for the problem of selecting an ARP. Let

First stage, regression slope

Recall from the discussion above that we regard an ARP as a portfolio that achieves a constant return per time period. This implies that when we regress the return from the portfolio against time we should have a regression slope of zero.

Obviously a regression slope of zero may not be attainable, but we can adopt an optimisation framework and choose a portfolio that minimises |regression slope| (i.e. has a slope that is as close to zero as possible).

We make the usual approximation assumption that portfolio return is a weighted sum of individual asset returns, i.e. that at time t portfolio return is given by $\sum_{i=1}^{N} w_i r_{it}$. Then, since (see notation in Section 3.2.2) $\hat{\beta}_i$ is the least-squares regression slope for asset i when the returns r_{it} from asset i are regressed against time, we have that the regression slope for the ARP when its returns ($\sum_{i=1}^{N} w_i r_{it}$) are regressed against time is given by $\sum_{i=1}^{N} w_i \hat{\beta}_i$.

Hence our first stage objective is:

$$\text{minimise } |\sum_{i=1}^{N} w_i \hat{\beta}_i| \tag{3.11}$$

Although this is nonlinear it can be linearised in a standard way. Introduce $E \geq 0$ and then:

$$\text{minimise } E \tag{3.12}$$

subject to (3.1)-(3.10) and:

$$E \geq \sum_{i=1}^{N} w_i \hat{\beta}_i \tag{3.13}$$

$$E \geq -(\sum_{i=1}^{N} w_i \hat{\beta}_i) \tag{3.14}$$

$$E \geq 0 \tag{3.15}$$

Second stage, regression intercept

Let E^* be the optimal value for E as found at the first stage above. Then having achieved the minimal absolute value for regression slope at the first stage we, at the second stage, maximise the regression intercept. Ideally if the regression slope is zero this stage maximises the (constant) return we achieve in each and every time period. This relates to the point discussed above in that we allow the return achieved by the ARP to be determined as a result of optimisation, it is not pre-specified.

Since (see notation in Section 3.2.2) $\hat{\alpha}_i$ is the least-squares regression intercept for asset i when the returns r_{it} from asset i are regressed against time we have that the regression intercept for the ARP when its returns ($\sum_{i=1}^{N} w_i r_{it}$) are regressed against time is given by $\sum_{i=1}^{N} w_i \hat{\alpha}_i$.

Hence at the second stage our optimisation is:

$$\text{maximise} \sum_{i=1}^{N} w_i \hat{\alpha}_i \tag{3.16}$$

subject to (3.1)-(3.10), (3.13)-(3.15) and:

$$E^* - \tau \leq E \leq E^* + \tau \tag{3.17}$$

In this stage we maximise the regression intercept whilst maintaining the regression slope at its (minimised) value as found at the first stage. Here τ is a small positive constant to cope with the fact that we cannot get exact accuracy from any numeric software and so imposing $E = E^*$ as a constraint is too restrictive. In the computational results reported later below we used $\tau = 0.0001$.

Third stage, transaction cost

Let Δ^* be the optimal value for the regression intercept $\sum_{i=1}^{N} w_i \hat{\alpha}_i$ as achieved at the second stage above. In the third stage we minimise the transaction cost associated with achieving the optimal values for regression slope and intercept as at the first two stages.

Hence at the third stage our optimisation is:

$$\text{minimise} \sum_{i=1}^{N} G_i \tag{3.18}$$

subject to (3.1)-(3.10), (3.13)-(3.15), (3.17) and:

$$\Delta^* - \tau \leq \sum_{i=1}^{N} w_i \hat{\alpha}_i \leq \Delta^* + \tau \tag{3.19}$$

3.2.5 Unavoidable transaction cost

Early computational results indicated that the three-stage model presented above encountered difficulties in terms of satisfying the transaction cost constraint (Equation (3.6)) in certain circumstances. As will become apparent in the computational results presented later below we use our model in a repeated fashion over time, rebalancing our ARP at various points in time. When we decide to make a rebalance of our current portfolio $[X_i]$ some transaction cost may be ***unavoidable***. For instance if we are holding an asset then its price might have risen since we bought it such that we are now in breach of the upper proportion limit for that asset (Equation (3.2)). In such cases we must sell some of the asset to ensure that the holding is within the limit. In another case we may have fallen below the lower proportion limit, here we can either buy some of the asset to reach the lower limit, or sell all of it.

Given our current holding X_i let $W_i = X_i V_{iT}/C$ represent the proportion invested in asset i in the current portfolio. If $X_i > 0$, the proportion limits for asset i are $\varepsilon_i \leq W_i \leq \delta_i$.

If $W_i > \delta_i$, we need to sell some of the asset. The maximum amount we can hold and satisfy the proportion limit is given by $\delta_i C/V_{iT}$. Hence we sell $[X_i - \delta_i C/V_{iT}]$ and this will incur a transaction cost of $f_i^s[X_i - \delta_i C/V_{iT}]V_{iT} = f_i^s[X_i V_{iT} - \delta_i C]$.

If $W_i < \varepsilon_i$, we could conceivably sell the asset. However, in this process we would be forcing another asset into the portfolio to satisfy Equation (3.3). Therefore, we shall adopt the view here that we need to trade to bring the asset back up to the minimum proportion. The minimum amount we can hold in order to satisfy the proportion limit is given by $\varepsilon_i C/V_{iT}$. We then buy $[\varepsilon_i C/V_{iT} - X_i]$ and this will incur a transaction cost of $f_i^b[\varepsilon_i C/V_{iT} - X_i]V_{iT} = f_i^b[\varepsilon_i C - X_i V_{iT}]$.

Hence, the unavoidable transaction cost we face under our proportion constraints is G^* given by:

$$G^* = \sum_{i=1,\ X_i>0,\ W_i>\delta_i}^{N} f_i^s[X_i V_{iT} - \delta_i C] + \sum_{i=1,\ X_i>0,\ W_i<\varepsilon_i}^{N} f_i^b[\varepsilon_i C - X_i V_{iT}] \tag{3.20}$$

Then Equation (3.6) becomes:

$$\sum_{i=1}^{N} G_i \leq \gamma C + G^* \tag{3.21}$$

Here we have redefined γ to be the limit $(0 \leq \gamma \leq 1)$ on the proportion of C that can be consumed by ***avoidable*** transaction cost. We accordingly replace Equation (3.6) with Equation (3.21) in the formulation presented above.

3.2.6 Summary

We have in this section set out our three-stage mixed-integer zero-one program for deciding an ARP. We refer to the model presented above as the ARP based on the regression of Return against Time, ***ARP-RT***.

Our approach however is flexible and we are able to extend it to the problem of designing portfolios with differing characteristics. This we do in the next section below.

3.3 Extensions

There are two extensions/amendments to the formulation presented above that can be made:

- the first extension relates to designing a portfolio that gives a constant excess return, return over and above the market index. This is therefore a model for a relative return portfolio, enhanced indexation (return above the index).
- the second extension relates to designing a portfolio with mixed characteristics, so a portfolio that is a mix of absolute and relative return.

We deal with each of these extensions below.

3.3.1 Enhanced indexation (relative return) portfolio

One argument that can be advanced against ARPs is that in good times (when the market/index is rising) it is a poor investment strategy to aim for an absolute return. Rather one should aim to do better than the index and produce a relative return portfolio, enhanced indexation. Due to the flexibility of our model we can easily amend it to produce portfolios that are designed to out-perform an index. For simplicity we shall continue to call the portfolio produced an ARP, rather than an enhanced indexation portfolio.

Suppose we regress the excess return of our chosen ARP (so return over and above index return) against time over the period we are considering. Our approach to enhanced indexation is to say that, ideally, this regression would have a slope of zero. This equates to a portfolio that (over time) has a constant (expected) excess return per time period (that return being given by the regression intercept).

Now (see notation above) $\hat{A}$ and $\hat{B}$ are the least-squares regression intercept and slope when the returns from the index (R_t) are regressed against time. Hence we have that when the excess return from the ARP ($\sum_{i=1}^{N} w_i r_{it} - R_t$) is regressed against time it will have regression intercept ($\sum_{i=1}^{N} w_i \hat{\alpha}_i - \hat{A}$) and regression slope ($\sum_{i=1}^{N} w_i \hat{\beta}_i - \hat{B}$).

Our first stage optimisation for the excess return regression slope is to try and achieve a regression slope that is, in absolute value terms, as close to zero as possible. This is therefore minimise $|\sum_{i=1}^{N} w_i\hat{\beta}_i - \hat{B}|$, which can be linearised to:

$$\text{minimise } E \tag{3.22}$$

subject to (3.1)-(3.5), (3.7)-(3.10), (3.15), (3.21) and:

$$E \geq \sum_{i=1}^{N} w_i\hat{\beta}_i - \hat{B} \tag{3.23}$$

$$E \geq -(\sum_{i=1}^{N} w_i\hat{\beta}_i - \hat{B}) \tag{3.24}$$

Our second stage optimisation for the regression intercept, to try and achieve a regression intercept that is as large as possible, has an objective function that is maximise $\sum_{i=1}^{N} w_i\hat{\alpha}_i - \hat{A}$. In this objective $\hat{A}$ is a constant and so can be ignored. Hence we have that the second stage objective here is precisely the same as the second stage objective given above, Equation (3.16), where this objective is optimised subject to (3.1)-(3.5), (3.7)-(3.10), (3.15), (3.17), (3.21) and (3.23)-(3.24).

The third stage follows in a similar fashion as for the ARP-RT model given above. Here the objective is to optimise Equation (3.18) subject to (3.1)-(3.5), (3.7)-(3.10), (3.15), (3.17), (3.19), (3.21) and (3.23)-(3.24).

We refer to the model presented here as the ARP based on the regression of Excess Return against Time, ***ARP-ERT***.

3.3.2 Mixed portfolio

In ARP-RT as presented above we have a pure absolute return model, whereas in ARP-ERT as presented above we have a pure enhanced indexation (relative return) model. It is possible to combine both models to produce portfolios with mixed characteristics - so a combined absolute return/relative return portfolio. Again for simplicity we shall continue to call the portfolio produced an ARP.

Let $\lambda \geq 0$ represent the weight that we attach to relative return as compared to absolute return. In the first stage optimisation for the regression slope we minimise $\max[|\sum_{i=1}^{N} w_i\hat{\beta}_i|, \lambda|\sum_{i=1}^{N} w_i\hat{\beta}_i - \hat{B}|]$, so minimise the maximum absolute value of both regression slopes (for the regressions of return against time and excess return against time) considered individually. Here we have introduced λ as a weighting for the regression slope associated with the relative return component of the objective. Again this is nonlinear but can be linearised as:

$$\text{minimise } E \tag{3.25}$$

subject to (3.1)-(3.5), (3.7)-(3.10), (3.13)-(3.15), (3.21) and:

$$E \geq \lambda(\sum_{i=1}^{N} w_i \hat{\beta}_i - \hat{B}) \tag{3.26}$$

$$E \geq -\lambda(\sum_{i=1}^{N} w_i \hat{\beta}_i - \hat{B}) \tag{3.27}$$

During the second stage optimisation, we maximise $[\min \sum_{i=1}^{N} w_i \hat{\alpha}_i, \lambda(\sum_{i=1}^{N} w_i \hat{\alpha}_i - \hat{A})]$, so maximise the minimum value of both regression intercepts considered individually. Although this is a nonlinear objective as $\hat{A}$ and λ are both constants we can simplify it to maximise $\sum_{i=1}^{N} w_i \hat{\alpha}_i$. Hence we have that the second stage objective here is precisely the same as the second stage objective given above, Equation (3.16), where this objective is optimised subject to subject to (3.1)-(3.5), (3.7)-(3.10), (3.13)-(3.15), (3.17), (3.21) and (3.26)-(3.27).

The third stage follows in a similar fashion as for the ARP-RT model given above. Here the objective is to optimise Equation (3.18) subject to (3.1)-(3.5), (3.7)-(3.10), (3.13)-(3.15), (3.17), (3.19), (3.21) and (3.26)-(3.27).

We refer to the model presented here as the ARP based on the regression of Return and Excess Return against Time, ***ARP-RERT***.

3.4 Computational results

In this section we present computational results for our three models, ARP-RT, ARP-ERT and ARP-RERT. We used an Intel Core2 Duo CPU E8500 @ 3.16GHz with 4GB of RAM with Linux as the operating system. The code was written in C++ and CPLEX 12.1 (CPLEX Optimizer (2013)) was used as the mixed-integer solver. Detailed computation times are not given below since the time needed to solve the cases considered was effectively insignificant. For the worst case encountered computation time was approximately $17s$. On average the computation time over all cases considered below was approximately $1.4s$.

3.4.1 Data and methodology

In our computational experimentation we used real-world historical weekly data taken from the universe of assets defined by the S&P (Standard and Poor's) Global 1200 index and subindices over the period January 1999 to September 2006 (400 weeks of data).

This index and its various subindices can be seen in Figure 3.1. This choice of universe ensures coverage of the world's major equity markets and means that we are dealing with liquid assets. This data set was previously used in Meade & Beasley (2011) and has been manually adjusted to account for changes in index composition. This means that our models use no more data than was available at the time, removing susceptibility to the influence of survivor bias. Note here that the S&P World ex-US 700 index, although shown in Figure 3.1, was not used since it was only assembled in 2004.

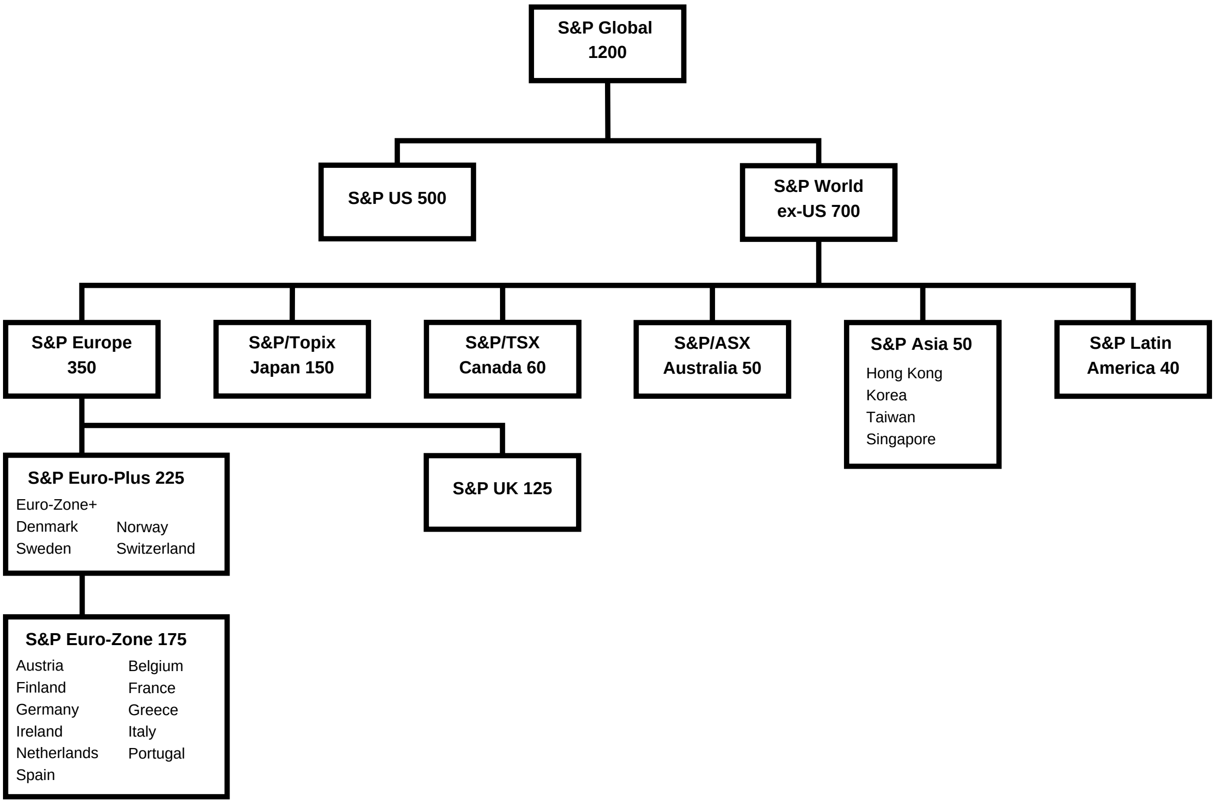

Figure 3.1: The structure of the S&P Global 1200 index, its components and the nations covered

The methodology we adopt is successive periodic rebalancing over time. We start from the beginning of our data set. We decide a portfolio using data taken from an in-sample period corresponding to the first h weeks. This portfolio is then held unchanged for an out-of-sample period of H weeks. We then rebalance (change) our portfolio, but now using the most recent h weeks as in-sample data. The decided portfolio is then held unchanged for an out-of-sample period of H weeks, and the process repeats until we have exhausted all of the data.

To illustrate this process suppose $h = 6$ and $H = 13$. The first step is portfolio selection based on the in-sample time period $[0,h]$=$[0,6]$. This time period contains 7 asset prices and so has $h = 6$ asset returns. We then evaluate the selected portfolio over the time period [6,6+H]=[6,19]. This time period gives us $H = 13$ out-of-sample returns, the first return being the return from time period 6 to time period 7, then from 7 to 8, etc. The next periods are [13,19] in-sample (containing 7 asset prices and $h = 6$ asset returns) and [19,32] out-of-sample, giving another 13 out-of-sample returns; then [26,32] in-sample and [32,45] out-of-sample; and so on until the data is exhausted. Once the data has been exhausted we have a time series of portfolio return values for out-of-sample performance, here from time period 7 (the first out-of-sample return value associated with the return from period 6 to period 7) until the end of the data (in other words we amalgamate together all out-of-sample return values) and this can be evaluated to see whether (or not) the cumulative effect of the decided portfolios has been favourable out-of-sample.

3.4.2 ARP evaluation and parameters

This evaluation of the out-of-sample return series for the ARP is a two-step procedure. For simplicity we describe the procedure for ARP-RT, the regression against time, but the procedure for ARP-ERT and ARP-RERT is similar. Recall that the logic behind ARP-RT is that we are seeking a portfolio that has a regression slope, when returns are regressed against time, of zero. The first step in our out-of-sample evaluation procedure therefore consists of performing a regression of out-of-sample returns against time and asking the question: is the regression slope zero?

Here clearly hypothesis testing plays a role. The null hypothesis is H_0:regression slope is zero versus the alternative hypothesis H_1:regression slope is different from zero, so a two-sided hypothesis test. Usually in hypothesis testing we are interested in rejecting the null hypothesis and so set a significance level of (typically) 5% or 1%. The significance level is the probability of rejecting H_0 when it is true. Here the situation is different. Here we are interested in not rejecting the null hypothesis and so we set a higher significance level (to make it easier to reject the null hypothesis). In the results presented below we use significance levels of 10% and higher.

If we reject the null hypothesis (at the specified significance level) then the portfolios produced by ARP-RT via successive rebalancing have (taken together) failed, in that we have statistical evidence that the out-of-sample regression slope is not zero. We refer to this hypothesis test as the ***slope test*** and passing the test (accepting H_0) as a ***slope test success***.

If we have a slope test success then the second stage logic for ARP-RT was to maximise the regression intercept. Although it was never explicit in our second stage optimisation clearly we would hope that the regression intercept would be positive (so a positive return). Hence the second step in our out-of-sample evaluation procedure is to look at the average out-of-sample return and conduct the one-sided hypothesis test H_0:average out-of-sample return is zero versus the alternative hypothesis H_1:average out-of-sample return is greater than zero. Here, since we are interested in rejecting H_0, we can judge the significance of the results obtained in a standard way, using p-values, as will be seen in the tables below.

On a technical note the reason why we conduct a hypothesis test on the average out-of-sample return (rather than the out-of-sample regression intercept) is that the two are equivalent given that slope test success means that the regression slope can be taken to be zero. Conditional on this information the standard least-squares regression equation for the regression intercept would give a regression intercept equal to the average out-of-sample return.

With regard to parameter values we examine a variety of values for the in-sample period h, and the out-of-sample period H, as will become apparent in the tables presented below. We set $C = 1000000$ corresponding to an initial investment of US$1million. For ARP-RERT we set the weighting parameter $\lambda = 1$, so weighting the absolute and relative return components in ARP-RERT equally. Although the models we have presented consider transaction cost, in the initial results below we allow trading to be free, so no cost associated with transactions, equivalently $f_i^b = f_i^s = 0$ for $i = 1, \ldots, N$. For the later results presented below we do impose a cost associated with transactions and a transaction cost limit.

We used $K = 0.8N$ as the number of assets to be in the portfolio. We set $\epsilon_i = 0.25/K$ and $\delta_i = 2/K$ for $i = 1, \ldots, N$ as the proportion limits for each asset in the portfolio. This ensures that the portfolios selected are not dominated by a small number of assets. Here the upper proportion limit is the same as that used previously in Meade & Beasley (2011). The lower proportion limit ensures that if an asset in present in the portfolio it is present at a reasonable level (having regard to the upper proportion limit already defined).

3.4.3 Results, zero transaction cost

In this section we give results for our three models ARP-RT, ARP-ERT and ARP-RERT at zero transaction cost. We first give results associated with a specific in-sample period of $h = 6$ weeks. We then go on to show how the results change as we change the in-sample period h. Finally we give Sharpe ratios for our chosen in-sample period.

In-sample period $h = 6$

In this section we give results for our three models ARP-RT, ARP-ERT and ARP-RERT for a specific in-sample period at zero transaction cost.

Table 3.1: Out-of-sample returns and excess returns for each model

Instance	H	ARP-RT		ARP-ERT		ARP-RERT		
		return	excess return	return	excess return	return	excess return	
S&P Latin America 40	4	0.00335$_{++}$	−0.00014	0.00308$_{++}$	−0.00041	0.00303$_{++}$	−0.00046	
	13	0.00271$_{+}$	−0.00078	0.00280$_{+}$	−0.00069	0.00281$_{+}$	−0.00068	
	26	0.00330$_{++}$	−0.00019	0.00336$_{++}$	−0.00013	0.00341$_{++}$	−0.00007	
	52	0.00285$_{+}$	−0.00064	0.00299$_{++}$	−0.00050	0.00292$_{+}$	−0.00057	
S&P Asia 50	4	0.00241$_{++}$	−0.00005	0.00233$_{++}$	−0.00013	0.00244$_{++}$	−0.00002	
	13	0.00353$_{++}$	0.00107$_{+}$	0.00351$_{++}$	0.00105	0.00337$_{++}$	0.00091	
	26	0.00349$_{+++}$	0.00103	0.00349$_{+++}$	0.00103	0.00350$_{+++}$	0.00104	
	52	0.00389$_{+++}$	0.00143$_{++}$	0.00374$_{+++}$	0.00128$_{+}$	0.00399$_{+++}$	0.00153$_{++}$	
S&P ASX 50	4	0.00302$_{+++}$	0.00120$_{++}$	0.00296$_{+++}$	0.00114$_{++}$	0.00296$_{+++}$	0.00113$_{++}$	
	13	0.00383$_{+++}$	(0.00200$_{+++}$)	0.00336$_{+++}$	0.00154$_{+++}$	0.00330$_{+++}$	0.00147$_{+++}$	
	26	0.00376$_{+++}$	0.00193$_{+++}$	0.00353$_{+++}$	0.00171$_{+++}$	0.00334$_{+++}$	0.00151$_{+++}$	
	52	0.00331$_{+++}$	0.00149$_{+++}$	0.00328$_{+++}$	0.00145$_{+++}$	0.00333$_{+++}$	0.00151$_{+++}$	
S&P TSX 60	4	0.00358$_{+++}$	0.00128	0.00343$_{+++}$	0.00113	0.00332$_{+++}$	0.00102	
	13	0.00381$_{+++}$	0.00151$_{+}$	0.00398$_{+++}$	0.00168$_{++}$	0.00405$_{+++}$	0.00175$_{++}$	
	26	0.00431$_{+++}$	0.00200$_{++}$	0.00451$_{+++}$	0.00221$_{+++}$	0.00450$_{+++}$	0.00220$_{+++}$	
	52	0.00279$_{++}$	0.00049	0.00413$_{+++}$	0.00183$_{++}$	0.00390$_{+++}$	0.00160$_{++}$	
S&P UK 125	4	0.00119	0.00078	0.00120	0.00079	0.00119	0.00078	
	13	0.00159	0.00117$_{+}$	0.00157	0.00116$_{+}$	0.00158	0.00117$_{+}$	
	26	0.00165$_{+}$	0.00124$_{++}$	0.00163$_{+}$	0.00122$_{++}$	0.00161	0.00119$_{++}$	
	52	0.00205$_{++}$	(0.00164$_{+++}$)	0.00207$_{++}$	(0.00166$_{+++}$)	0.00206$_{++}$	(0.00165$_{+++}$)	
S&P Topix 150	4	0.00132	0.00029	0.00135	0.00032	0.00144	0.00041	
	13	0.00260$_{+}$	(0.00157$_{+++}$)	0.00258$_{+}$	0.00154$_{++}$	0.00259$_{+}$	0.00156$_{+++}$	
	26	0.00283$_{++}$	0.00180$_{+++}$	0.00276$_{++}$	0.00173$_{+++}$	0.00282$_{++}$	0.00179$_{+++}$	
	52	0.00365$_{++}$	(0.00262$_{+++}$)	0.00368$_{++}$	(0.00265$_{+++}$)	0.00374$_{++}$	(0.00271$_{+++}$)	
S&P Euro Zone 175	4	0.00162$_{+}$	0.00082	0.00177$_{+}$	0.00097$_{+}$	0.00171$_{+}$	0.00091	
	13	0.00197$_{+}$	0.00117$_{++}$	0.00195$_{+}$	0.00115$_{++}$	0.00191$_{+}$	0.00111$_{++}$	
	26	0.00197$_{+}$	0.00117$_{++}$	0.00201$_{++}$	0.00120$_{++}$	0.00193$_{+}$	0.00113$_{++}$	
	52	0.00212$_{++}$	0.00131$_{++}$	0.00216$_{++}$	0.00136$_{++}$	0.00213$_{++}$	0.00132$_{++}$	
S&P Euro Plus 225	4	0.00155$_{+}$	0.00070	0.00157$_{+}$	0.00072	0.00164$_{+}$	0.00079	
	13	0.00172$_{+}$	0.00086$_{+}$	0.00181$_{+}$	0.00096$_{+}$	0.00175$_{+}$	0.00090$_{+}$	
	26	0.00189$_{+}$	0.00104$_{+}$	0.00200$_{++}$	0.00115$_{++}$	0.00197$_{+}$	0.00111$_{++}$	
	52	0.00214$_{++}$	0.00128$_{++}$	0.00210$_{++}$	0.00124$_{++}$	0.00214$_{++}$	0.00128$_{++}$	
S&P Europe 350	4	0.00150$_{+}$	0.00081$_{+}$	0.00144$_{+}$	0.00075$_{+}$	0.00142	0.00074	
	13	0.00170$_{+}$	0.00102$_{++}$	0.00177$_{+}$	0.00108$_{++}$	0.00171$_{+}$	0.00102$_{++}$	
	26	0.00195$_{+}$	0.00126$_{+++}$	0.00196$_{++}$	0.00127$_{+++}$	0.00192$_{+}$	0.00123$_{++}$	
	52	0.00203$_{++}$	0.00134$_{+++}$	0.00199$_{++}$	0.00130$_{+++}$	0.00199$_{++}$	0.00130$_{+++}$	
S&P US 500	4	0.00109	0.00094$_{+}$	0.00120	0.00106$_{++}$	0.00108	0.00093$_{+}$	
	13	0.00129	0.00115$_{++}$	0.00134	0.00120$_{++}$	0.00128	0.00114$_{++}$	
	26	0.00189$_{+}$	0.00175$_{+++}$	0.00188$_{+}$	0.00174$_{+++}$	0.00189$_{+}$	0.00175$_{+++}$	
	52	0.00154$_{+}$	0.00140$_{+++}$	0.00158$_{+}$	0.00144$_{+++}$	0.00158$_{+}$	0.00144$_{+++}$	
S&P Global 1200	4	0.00165$_{+}$	0.00114$_{++}$	0.00165$_{+}$	0.00114$_{++}$	0.00162$_{+}$	0.00110$_{++}$	
	13	0.00193$_{++}$	0.00142$_{+++}$	0.00187$_{++}$	0.00136$_{+++}$	0.00191$_{++}$	0.00139$_{+++}$	
	26	0.00229$_{++}$	0.00178$_{+++}$	0.00235$_{++}$	0.00184$_{+++}$	0.00233$_{++}$	0.00182$_{+++}$	
	52	0.00244$_{++}$	(0.00193$_{+++}$)	0.00252$_{+++}$	(0.00200$_{+++}$)	0.00244$_{++}$	(0.00193$_{+++}$)	
# of slope test successes		44	39	44	41	44	41	**253**
# of entries	(+)	16	7	12	5	14	3	**57**
# of entries	(++)	14	11	16	14	13	14	**82**
# of entries	(+++)	9	9	11	11	10	11	**61**
# total		39	27	39	30	37	28	**200**

Note: Under the one-sided alternative hypothesis that the return / excess return is greater than zero, + means $p < 10\%$, ++ means $p < 5\%$, +++ means $p < 1\%$

The results in Table 3.1 compare each model (ARP-RT, ARP-ERT and ARP-RERT) for different values of the out-of-sample holding period H. The ***return*** columns are the average out-of-sample weekly continuous time returns and the ***excess return*** columns represent the average out-of-sample excess over the average index weekly continuous time return. This table uses an in-sample period h of 6 weeks.

For the slope test we used a significance level of 10%. An entry not enclosed in brackets in Table 3.1 indicates a slope test success at this significance level, an entry in brackets indicates a slope test failure at this significance level. In this table slope test failures are only seen for the excess return columns, meaning that an ARP for which excess return that was, in regression terms, statistically independent of time, could not be achieved for that particular set {model, instance, H}. The total number of slope test successes can be seen near the foot of Table 3.1.

Recall that given a slope test success we test the average out-of-sample return, and success here is indicated by the p-values shown using plus subscripts. For example, consider the entry of 0.00229_{++} in the return column for ARP-RT for the S&P Global 1200 with an out-of-sample period of $H = 26$ weeks. As this is not enclosed in brackets there was a slope test success for this case. The average out-of-sample weekly return is 0.00229 (this is a continuous time fractional return, not a percentage) which corresponds to a yearly percentage of $100(\exp(0.00229 \times 52) - 1) = 12.6\%$. The double plus subscript means that this is associated with a p-value of less than 5%. In other words there is a probability of this result arising by chance of less than 5% if the true underlying average out-of-sample return was actually zero. The total number of entries for each p-value (but excluding slope test failures, so entries in brackets) can be seen at the foot of Table 3.1.

Associated with the entry of 0.00229_{++} in the return column for ARP-RT for the S&P Global 1200 with $H = 26$ we also see an entry of 0.00178_{+++} in the excess return column. This corresponds to an evaluation of out-of-sample portfolio excess returns (returns over and above the index). The average out-of-sample weekly excess return is 0.00178 which corresponds to a yearly percentage of $100(\exp(0.00178 \times 52) - 1) = 9.7\%$. The triple plus subscript means that this is associated with a p-value of less than 1%.

Figure 3.2 shows the (normalised) value of the ARP-RT portfolio, as well as the index value, over the out-of-sample period. Notice in particular how the portfolio substantially out-performs the index over such a long time period. Moreover this figure exhibits the behaviour we would expect if we were indeed achieving constant return per time period, namely growth in portfolio value appearing to have a linear relationship with time.

Considering Table 3.1 we can observe that for all cases {model, instance, H}, the average returns obtained by our models were strictly positive. Some of the average excess returns however were negative. Although not all of the cases passed the slope test the

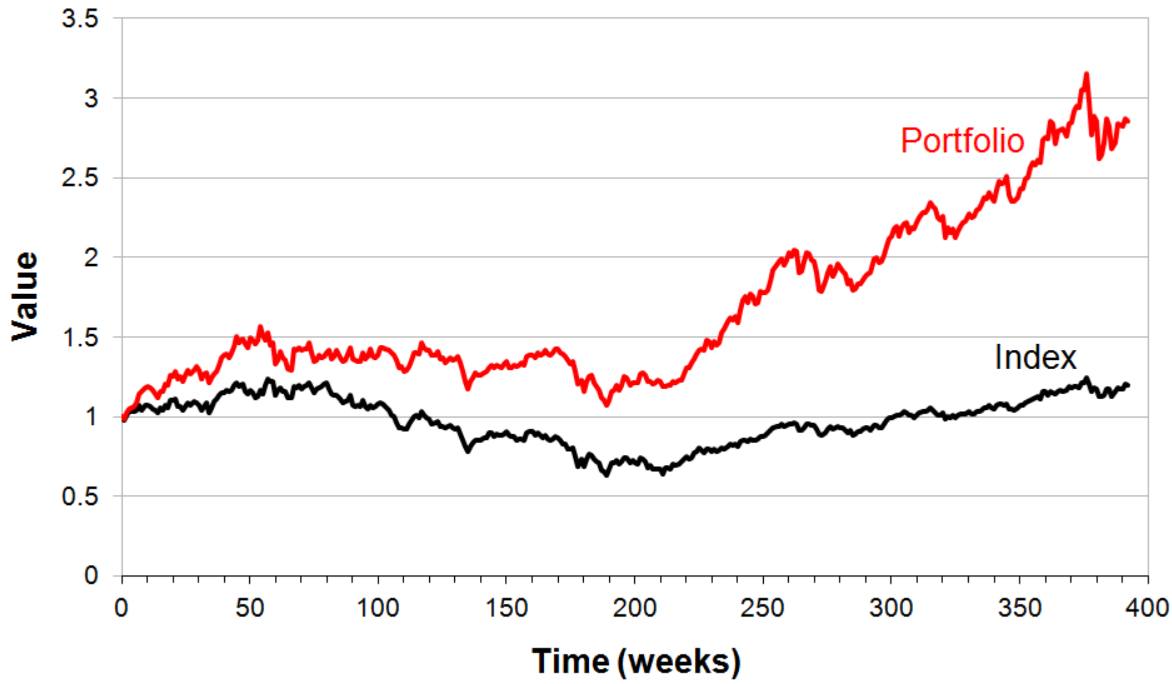

Figure 3.2: Out-of-sample portfolio and index value, S&P Global 1200, ARP-RT, $H = 26$

vast majority (253 out of 264 cases) did. Clearly not every positive return/excess return is (statistically) judged to be significant (greater than zero). However, over all the 264 entries shown in Table 3.1, 61 that passed the slope test were significant at the 1% level and a further 82 were significant at the 5% level.

In Table 3.1 we have used a significance level for the slope test of 10%. Recall that the significance level is the probability of rejecting the null hypothesis (H_0:regression slope is zero) when it is true. If we increase the significance level then we are making it harder to pass the slope test since we are increasing the probability of rejecting it when it is true. Table 3.2 shows summary results as we increase the significance level from 10% upward. The returns and excess returns for each case are omitted from this table since they are the same as in Table 3.1.

Considering Table 3.2 we can see that, as expected, we have a decrease in the number of slope test successes as the significance level increases. However the effect is not especially pronounced, neither is there a large effect in terms of a decrease in the number of entries for the higher p-values (1% and 5%) until we reach a significance level of 50%. Overall Table 3.2 does not indicate that the results in Table 3.1 are especially sensitive to the slope test significance level adopted until we reach very high levels of significance.

Significance Level		ARP-RT		ARP-ERT		ARP-RERT		
		return	excess return	return	excess return	return	excess return	
10%	# of slope test successes	44	39	44	41	44	41	**253**
	# of entries (+)	16	7	12	5	14	3	**57**
	# of entries (++)	14	11	16	14	13	14	**82**
	# of entries (+++)	9	9	11	11	10	11	**61**
	# total	39	27	39	30	37	28	**200**
20%	# of slope test successes	40	35	40	35	41	35	**226**
	# of entries (+)	13	7	9	5	11	3	**48**
	# of entries (++)	13	11	15	13	13	14	**79**
	# of entries (+++)	9	5	11	7	10	5	**47**
	# total	35	23	35	25	34	22	**174**
30%	# of slope test successes	33	31	34	30	34	29	**191**
	# of entries (+)	7	6	6	5	6	3	**33**
	# of entries (++)	13	11	13	12	13	13	**75**
	# of entries (+++)	9	3	11	4	10	2	**39**
	# total	29	20	30	21	29	18	**147**
50%	# of slope test successes	21	16	20	18	19	18	**112**
	# of entries (+)	5	6	4	4	4	3	**26**
	# of entries (++)	8	5	7	6	7	6	**39**
	# of entries (+++)	5	0	6	1	5	1	**18**
	# total	18	11	17	11	16	10	**83**

Note: Under the one-sided alternative hypothesis that the return / excess return is greater than zero, + means $p < 10\%$, ++ means $p < 5\%$, +++ means $p < 1\%$

Table 3.2: Summary table for the slope test at varying significance levels

Varying the in-sample period h

Above we have considered a specific in-sample period of $h = 6$ weeks. To illustrate how the results change as we vary the in-sample period we have Table 3.3. In that table we show, for varying combinations of in-sample period h and out-of-sample period H, the average returns and excess returns (averaged over all instances and all three models combined). The entries in Table 3.3 for $h = 6$ are the averages of the appropriate results seen in Table 3.1.

Taking the $h = 6$, $H = 4$ entry in Table 3.3 we have that over all three models the average out-of-sample weekly return was 0.00200. The standard deviation associated with these returns was 0.00082. With three models and eleven data instances the maximum number of slope test successes is 33, and here in fact we did have 33 slope test successes. For the 33 returns for which the slope test was a success the number that were significant at each level can be seen (so for example 6 of these returns had a p-value less than 1%).

Considering Table 3.3 we can see that, judging from the averages at least, our models generate positive returns and excess returns for all combinations of h and H. ***Broadly the results in Table 3.3 indicate that, for the particular instances we have***

		$H = 4$		$H = 13$		$H = 26$		$H = 52$	
		return	excess return	return	excess return	return	excess return	return	excess return
$h = 6$	Average	0.00200	0.00068	0.00241	0.00109	0.00267	0.00135	0.00271	0.00138
	Standard deviation	0.00082	0.00049	0.00088	0.00064	0.00089	0.00059	0.00080	0.00076
	# of slope test successes	33	33	33	31	33	33	33	24
	# of entries (+)	11	5	15	8	11	1	5	1
	# of entries (++)	6	7	6	12	12	10	19	10
	# of entries (+++)	6	0	6	6	9	16	9	9
	# total	23	12	27	26	32	27	33	20
$h = 13$	Average	0.00194	0.00077	0.00227	0.00110	0.00246	0.00129	0.00253	0.00136
	Standard deviation	0.00073	0.00082	0.00083	0.00055	0.00082	0.00068	0.00085	0.00081
	# of slope test successes	33	32	33	29	33	31	33	19
	# of entries (+)	13	15	13	3	10	3	8	0
	# of entries (++)	2	11	8	13	14	9	16	7
	# of entries (+++)	5	0	6	7	6	13	6	9
	# total	20	26	27	23	30	25	30	16
$h = 26$	Average	0.00198	0.00103	0.00192	0.00097	0.00247	0.00152	0.00221	0.00126
	Standard deviation	0.00081	0.00040	0.00078	0.00060	0.00074	0.00052	0.00067	0.00046
	# of slope test successes	21	32	28	33	30	27	32	29
	# of entries (+)	4	5	7	4	8	1	12	2
	# of entries (++)	3	8	3	9	13	8	6	14
	# of entries (+++)	2	7	4	9	6	13	5	8
	# total	9	20	14	22	27	22	23	24
$h = 39$	Average	0.00214	0.00107	0.00242	0.00135	0.00237	0.00130	0.00217	0.00110
	Standard deviation	0.00073	0.00048	0.00080	0.00059	0.00070	0.00056	0.00068	0.00046
	# of slope test successes	33	27	33	24	33	22	30	28
	# of entries (+)	13	9	10	1	11	1	13	2
	# of entries (++)	5	9	12	7	13	9	6	11
	# of entries (+++)	6	3	8	12	6	7	5	7
	# total	24	21	30	20	30	17	24	20
$h = 52$	Average	0.00185	0.00124	0.00163	0.00103	0.00160	0.00099	0.00166	0.00106
	Standard deviation	0.00089	0.00055	0.00085	0.00051	0.00093	0.00042	0.00087	0.00039
	# of slope test successes	16	26	9	31	9	30	10	33
	# of entries (+)	3	2	0	7	0	6	0	8
	# of entries (++)	5	14	2	11	3	10	3	13
	# of entries (+++)	3	5	4	6	3	4	3	8
	# total	11	21	6	24	6	20	6	29

Note: Under the one-sided alternative hypothesis that the return / excess return is greater than zero, + means $p < 10\%$, ++ means $p < 5\%$, +++ means $p < 1\%$

Table 3.3: Average return and excess return for each choice of h and H

examined, relatively short in-sample periods are sufficient to generate good out-of-sample performance for relatively long periods. We did investigate this in much greater detail in order to understand it better, since it did run counter to what we would have intuitively expected.

Recall that we must have $K = 0.8N$ assets in the portfolio, so it must contain a large number of assets and is very diversified. Each asset in the portfolio, in proportion terms, must lie between $0.25/K$ and $2/K$ of the portfolio, so it is relatively restricted. This means that at each rebalance there is limited freedom (for example in terms of the assets held in the portfolio, even if we have no transaction cost constraint). For this reason a relatively short in-sample period is sufficient to generate a 'good' portfolio.

Sharpe ratio

To further validate our models we also calculated the (yearly) Sharpe ratio for the out-of-sample returns and excess returns, as presented in Table 3.4. The Sharpe ratio (Sharpe (1966, 1975, 1994)) is the excess return (excess being defined here as portfolio return minus the risk-free rate) divided by the standard deviation of portfolio return. It is a numeric measure that captures the tradeoff between return achieved in excess of the risk-free rate and risk (as measured by variation in return). The higher the Sharpe ratio, the better. For comparison purposes the Sharpe ratio obtained if we had exactly tracked the index is also presented in Table 3.4. Here for a risk-free rate we used historic values for the 3 month rate for US Treasury Constant Maturities.

Taking ARP-RT and $H = 26$ for the S&P Global 1200 we have in Table 3.4 that the yearly Sharpe ratio if we had exactly tracked the index would have been -0.02988. Here the negative sign indicates that over the out-of-sample period the return from the index did not exceed the risk-free return. The portfolios produced by our ARP-RT model had an out-of-sample performance corresponding to a yearly Sharpe ratio of 0.58879. Here the (implicit) positive sign indicates that we did exceed the risk-free rate. Examining Table 3.4 we can see that the returns from all three models produce superior Sharpe ratios to the index Sharpe ratio except for the S&P Latin America 40.

3.4.4 Results, transaction cost

In this section we give results for our three models ARP-RT, ARP-ERT and ARP-RERT when we incur transaction costs associated with trading at each rebalance, and also have a constraint on total transaction costs incurred at each rebalance. We set the costs associated with buying and selling using $f_i^b = f_i^s = 0.005$ for $i = 1, \ldots, N$. This effectively corresponds to a round trip transaction cost of 1% (all costs associated with opening and closing a financial transaction, see Meade & Beasley (2011) for a discussion as to estimated round trip transaction costs). In terms of the limit on (avoidable) transaction cost, Equation (3.21), we used $\gamma = 0.005$, so we were prepared to sacrifice at most 0.5% of the portfolio value at each rebalance in avoidable transaction cost.

The results relating to transaction cost can be seen in Table 3.5. In that table we show the returns and excess returns for our three models ARP-RT, ARP-ERT and ARP-RERT for an in-sample period of $h = 6$. This table is therefore comparable with Table 3.1 which was for the same in-sample period but with zero transaction cost.

Table 3.4: Sharpe ratios

Instance	H	ARP-RT		ARP-ERT		ARP-RERT	
		return	excess return	return	excess return	return	excess return
S&P Latin America 40	4	0.54814	−0.31723	0.49899	−0.43870	0.48902	−0.46246
	13	0.42077	−0.63820	0.43241	−0.60572	0.44382	−0.58573
0.57367	26	0.55006	−0.37329	0.55762	−0.35105	0.57010	−0.32153
	52	0.46329	−0.64154	0.48446	−0.56104	0.46132	−0.59870
S&P Asia 50	4	0.46626	−0.29596	0.44614	−0.32858	0.47734	−0.27806
	13	0.69968	0.20531	0.70140	0.19388	0.66502	0.13755
0.40393	26	0.72171	0.18331	0.72727	0.18359	0.72411	0.18684
	52	0.77731	0.36969	0.73984	0.30042	0.80641	0.41248
S&P ASX 50	4	0.75555	0.35483	0.73632	0.32365	0.73440	0.32260
	13	0.90181	0.66102	0.84575	0.62110	0.82601	0.56338
0.36924	26	0.95936	0.83475	0.89995	0.70651	0.83852	0.59016
	52	0.81679	0.57352	0.82520	0.55187	0.84003	0.57728
S&P TSX 60	4	0.88054	0.23966	0.85148	0.18901	0.81042	0.15128
	13	0.84108	0.33567	0.87617	0.41516	0.91040	0.44844
0.42601	26	0.95019	0.54367	0.98430	0.64711	0.97811	0.64040
	52	0.59228	−0.04585	0.86594	0.48814	0.81822	0.38718
S&P UK 125	4	0.18647	0.09672	0.18956	0.10236	0.18528	0.09503
	13	0.27450	0.27175	0.27164	0.26767	0.27451	0.27084
−0.05901	26	0.29975	0.33301	0.29376	0.32259	0.28678	0.30820
	52	0.43000	0.56350	0.43373	0.57080	0.43086	0.56510
S&P Topix 150	4	0.16285	−0.16518	0.16990	−0.14980	0.19117	−0.10212
	13	0.45211	0.53166	0.44632	0.51637	0.44851	0.52475
0.09071	26	0.48896	0.69545	0.47483	0.65012	0.48919	0.68972
	52	0.62657	1.11612	0.62790	1.12033	0.64246	1.16351
S&P Euro Zone 175	4	0.31064	0.10888	0.35632	0.18406	0.33977	0.15432
	13	0.39486	0.31664	0.38810	0.30620	0.37573	0.28162
0.04990	26	0.41539	0.30717	0.42737	0.31908	0.40641	0.28372
	52	0.45564	0.38657	0.46831	0.41546	0.45733	0.39686
S&P Euro Plus 225	4	0.28783	0.05106	0.29441	0.06174	0.31621	0.10196
	13	0.31697	0.15168	0.34341	0.20920	0.32549	0.17314
0.06440	26	0.39215	0.24550	0.42553	0.30718	0.41451	0.29065
	52	0.46315	0.40547	0.44978	0.38057	0.46341	0.40525
S&P Europe 350	4	0.28814	0.12838	0.27063	0.09461	0.26376	0.08235
	13	0.32170	0.26912	0.34150	0.31137	0.32476	0.27465
0.02387	26	0.41209	0.45067	0.41481	0.45452	0.40342	0.42665
	52	0.44589	0.50971	0.43248	0.48203	0.43110	0.48468
S&P US 500	4	0.14820	0.20968	0.18361	0.28257	0.14515	0.20305
	13	0.19177	0.32109	0.20537	0.35349	0.18854	0.31137
−0.13630	26	0.36938	0.78238	0.36359	0.78757	0.36894	0.79760
	52	0.28684	0.68293	0.29736	0.72240	0.29799	0.71720
S&P Global 1200	4	0.37647	0.35234	0.37617	0.35309	0.36258	0.32797
	13	0.44116	0.55463	0.41912	0.51588	0.43362	0.55199
−0.02988	26	0.58879	0.91449	0.60957	0.96983	0.60348	0.94853
	52	0.62215	1.13480	0.64703	1.18624	0.62395	1.11203

Conceptually transaction costs can be viewed as a cost that is incurred (in full) now as the necessary price that we have to pay in order to change (rebalance) to a portfolio that will, based on in-sample data, perform better than our existing portfolio in the future.

Table 3.5: Out-of-sample returns and excess returns for each model, transaction cost case

Instance	H	ARP-RT		ARP-ERT		ARP-RERT		
		return	excess return	return	excess return	return	excess return	
S&P Latin America 40	4	0.00235_{+}	-0.00114	0.00261_{+}	-0.00088	0.00241_{+}	-0.00108	
	13	0.00288_{+}	-0.00061	0.00282_{+}	(-0.00067)	0.00315_{++}	-0.00034	
	26	0.00314_{++}	-0.00035	0.00327_{++}	-0.00022	0.00313_{++}	-0.00036	
	52	0.00269_{+}	-0.00080	0.00263_{+}	-0.00086	0.00264_{+}	-0.00085	
S&P Asia 50	4	0.00156	-0.00090	0.00172	-0.00074	0.00163	-0.00083	
	13	0.00336_{++}	0.00090	0.00353_{+++}	0.00107_{+}	0.00330_{++}	0.00084	
	26	0.00344_{++}	0.00098	0.00343_{++}	0.00097	0.00355_{++}	0.00109_{+}	
	52	0.00329_{++}	0.00083	0.00372_{+++}	0.00126_{+}	0.00370_{+++}	0.00124_{+}	
S&P ASX 50	4	0.00217_{++}	0.00035	0.00238_{++}	0.00056	0.00212_{++}	0.00029	
	13	0.00283_{+++}	0.00101_{++}	0.00284_{+++}	0.00102_{++}	0.00300_{+++}	0.00117_{++}	
	26	0.00349_{+++}	0.00166_{+++}	0.00338_{+++}	0.00155_{+++}	0.00341_{+++}	0.00159_{+++}	
	52	0.00313_{+++}	(0.00131_{++})	0.00299_{+++}	0.00116_{++}	0.00308_{+++}	0.00125_{++}	
S&P TSX 60	4	0.00225_{++}	-0.00005	0.00233_{++}	0.00003	0.00236_{++}	0.00006	
	13	0.00333_{+++}	0.00103	0.00378_{+++}	0.00148_{+}	0.00382_{+++}	0.00152_{+}	
	26	0.00425_{+++}	0.00195_{++}	0.00449_{+++}	0.00219_{+++}	0.00436_{+++}	0.00206_{++}	
	52	0.00257_{++}	0.00027	0.00391_{+++}	0.00160_{++}	0.00367_{+++}	0.00137_{+}	
S&P UK 125	4	0.00045	0.00004	0.00040	-0.00001	0.00045	0.00004	
	13	0.00133	0.00092	0.00133	0.00092	0.00132	0.00091	
	26	0.00185_{+}	0.00144_{++}	0.00187_{+}	0.00146_{++}	0.00174_{+}	0.00133_{++}	
	52	0.00190_{+}	0.00148_{++}	0.00202_{+}	(0.00160_{+++})	0.00190_{+}	0.00149_{++}	
S&P Topix 150	4	0.00033	-0.00070	0.00031	-0.00072	0.00029	-0.00075	
	13	0.00227_{+}	(0.00124_{++})	0.00227_{+}	(0.00124_{++})	0.00228_{+}	(0.00125_{++})	
	26	0.00282_{++}	(0.00179_{+++})	0.00275_{++}	(0.00172_{+++})	0.00276_{++}	(0.00173_{+++})	
	52	0.00346_{++}	(0.00243_{+++})	0.00350_{++}	(0.00247_{+++})	0.00350_{++}	(0.00247_{+++})	
S&P Euro Zone 175	4	0.00068	-0.00013	0.00051	-0.00030	0.00057	-0.00024	
	13	0.00190_{+}	0.00109_{++}	0.00165_{+}	0.00085_{+}	0.00160	0.00079	
	26	0.00168_{+}	0.00088	0.00165_{+}	0.00085	0.00169_{+}	0.00088_{+}	
	52	0.00210_{++}	0.00129_{++}	0.00214_{++}	0.00134_{++}	0.00212_{++}	0.00132_{++}	
S&P Euro Plus 225	4	0.00076	-0.00009	0.00075	-0.00011	0.00069	-0.00016	
	13	0.00162	0.00076	0.00159	0.00073	0.00161	0.00076	
	26	0.00173_{+}	0.00087_{+}	0.00174_{+}	0.00089_{+}	0.00176_{+}	0.00091_{+}	
	52	0.00209_{++}	0.00124_{++}	0.00206_{++}	0.00120_{++}	0.00200_{++}	0.00115_{++}	
S&P Europe 350	4	0.00046	-0.00022	0.00046	-0.00023	0.00045	-0.00024	
	13	0.00143	0.00074_{+}	0.00149	0.00080_{+}	0.00148	0.00079_{+}	
	26	0.00175_{+}	0.00106_{++}	0.00175_{+}	0.00106_{++}	0.00174_{+}	0.00105_{++}	
	52	0.00188_{+}	0.00119_{++}	0.00191_{+}	0.00122_{++}	0.00189_{+}	0.00120_{++}	
S&P US 500	4	0.00021	0.00006	0.00020	0.00006	0.00017	0.00002	
	13	0.00099	0.00084_{+}	0.00098	0.00083_{+}	0.00098	0.00084_{+}	
	26	0.00174_{+}	0.00160_{+++}	0.00181_{+}	0.00166_{+++}	0.00177_{+}	0.00163_{+++}	
	52	0.00164_{+}	0.00150_{+++}	0.00142	0.00128_{+++}	0.00162_{+}	0.00148_{+++}	
S&P Global 1200	4	0.00073	0.00022	0.00074	0.00023	0.00073	0.00022	
	13	0.00155_{+}	0.00104_{++}	0.00156_{+}	0.00105_{++}	0.00155_{+}	0.00104_{++}	
	26	0.00219_{++}	0.00167_{+++}	0.00222_{++}	0.00170_{+++}	0.00222_{++}	0.00171_{+++}	
	52	0.00225_{++}	(0.00174_{+++})	0.00233_{++}	(0.00182_{+++})	0.00230_{++}	(0.00179_{+++})	
# of slope test successes		44	39	44	38	44	40	**249**
# of entries	(+)	14	3	13	7	12	8	**57**
# of entries	(++)	13	10	10	9	12	10	**64**
# of entries	(+++)	5	4	8	5	7	4	**33**
# total		32	17	31	21	31	22	**154**

Note: Under the one-sided alternative hypothesis that the return / excess return is greater than zero, + means $p < 10\%$, ++ means $p < 5\%$, +++ means $p < 1\%$

As the results in Table 3.1 were produced at zero transaction cost (where trading was free) we would expect the transaction cost results in Table 3.5 to show reduced returns compared to Table 3.1. This is, in general, what we do observe. Comparing Table 3.1 and Table 3.5 we see a slight reduction in slope test successes (from 253 to 249, a reduction of 1.6%) but a large reduction in the number of significant p-values (e.g. for p-values at 5% and 1% we have a reduction from 82+61=143 to 64+33=97, a reduction of 32.2%).

The yearly Sharpe ratios associated with the results shown in Table 3.5 can be seen in Table 3.6. This table is therefore comparable with Table 3.4 which was for the same in-sample period but with zero transaction cost. Just as transaction costs reduce returns compared with the zero transaction cost case, so too they reduce Sharpe ratios. However from Table 3.6 it is clear that we still, for the majority of cases, produce Sharpe ratios superior to those exhibited by the indices themselves. One item of note that does come out more clearly from Table 3.6, as compared with Table 3.5, is the appearance of negative Sharpe ratios for a number of cases where $H = 4$. This out-of-sample holding period requires many more rebalances than the larger H values of $H = 13, 26, 52$ considered and this is reflected in the Sharpe ratios being much lower than those seen for $H = 4$ in Table 3.4.

3.4.5 Regression against time

Recall here that, as discussed above, we are ***defining*** an ARP as a portfolio that (ideally) achieves a constant return per time period. As such a regression of portfolio return against time is the appropriate regression to use. Regressing portfolio return against other factors (as in Carhart (1997); Fama & French (1993, 1996); Sharpe (1964)) would not satisfy the definition we have set out for an ARP. ***Although for any individual asset time may not be a significant explanatory variable of asset returns the key point here is whether, or not, by adopting a regression against time viewpoint we can construct portfolios that have the ARP characteristic of constant return that we desire.*** The results given above, verified by statistical hypothesis testing with respect to regression slope, did indicate that ARPs could be appropriately constructed.

As conducting a regression against time is not a usual procedure in terms of portfolio construction/analysis in this section we investigate our use of this regression in greater detail. Specifically we investigate:

- whether, when asset returns are regressed against time, we encounter significant linear regressions
- Type II errors, associated with falsely accepting the null hypothesis when testing a regression slope for significance

Table 3.6: Sharpe ratios, transaction cost case

Instance	H	ARP-RT		ARP-ERT		ARP-RERT	
		return	excess return	return	excess return	return	excess return
S&P Latin America 40	4	0.36401	−0.78790	0.42160	−0.67980	0.37372	−0.76015
	13	0.46669	−0.60838	0.44070	−0.66238	0.51552	−0.46234
0.57367	26	0.51999	−0.46570	0.56357	−0.39554	0.51377	−0.46413
	52	0.42956	−0.73906	0.41542	−0.76954	0.43015	−0.73546
S&P Asia 50	4	0.24548	−0.65453	0.28917	−0.57775	0.26470	−0.61931
	13	0.66898	0.13893	0.71006	0.21745	0.64760	0.10852
0.40393	26	0.67810	0.16853	0.68747	0.16732	0.70431	0.21849
	52	0.64188	0.09953	0.73945	0.29799	0.74756	0.28911
S&P ASX 50	4	0.48300	−0.15282	0.54909	−0.02830	0.46680	−0.18555
	13	0.67747	0.26282	0.68949	0.26926	0.72949	0.36631
0.36924	26	0.87267	0.65193	0.84662	0.58542	0.85669	0.61425
	52	0.76620	0.45378	0.72863	0.37100	0.76208	0.42069
S&P TSX 60	4	0.47571	−0.23275	0.49677	−0.20515	0.50857	−0.19436
	13	0.71717	0.16166	0.82195	0.34174	0.84143	0.35867
0.42601	26	0.95376	0.53594	0.98621	0.64183	0.96842	0.59187
	52	0.53046	−0.13263	0.81213	0.39353	0.76426	0.30066
S&P UK 125	4	−0.04781	−0.31443	−0.06542	−0.34453	−0.05026	−0.31947
	13	0.20617	0.15700	0.20492	0.15485	0.20230	0.15003
−0.05901	26	0.35592	0.44164	0.35805	0.45071	0.32341	0.38120
	52	0.37771	0.46618	0.41180	0.53250	0.37984	0.47041
S&P Topix 150	4	−0.06256	−0.66865	−0.06690	−0.67596	−0.07259	−0.68851
	13	0.37752	0.35699	0.37708	0.35490	0.37937	0.36083
0.09071	26	0.48753	0.69544	0.47238	0.65561	0.47648	0.65784
	52	0.58606	1.03832	0.59314	1.05328	0.59223	1.05381
S&P Euro Zone 175	4	0.02191	−0.37137	−0.02939	−0.45949	−0.01139	−0.42960
	13	0.37052	0.27736	0.29953	0.13742	0.28406	0.10570
0.04990	26	0.32775	0.14306	0.32066	0.12706	0.32916	0.14622
	52	0.45394	0.37035	0.46681	0.39436	0.46479	0.37541
S&P Euro Plus 225	4	0.04800	−0.38152	0.04319	−0.38971	0.02657	−0.41912
	13	0.28637	0.09397	0.27751	0.07662	0.28461	0.09035
0.06440	26	0.34096	0.15261	0.34493	0.16182	0.35099	0.17192
	52	0.44673	0.37319	0.44148	0.35061	0.42234	0.31788
S&P Europe 350	4	−0.04511	−0.51705	−0.04809	−0.52363	−0.04903	−0.52556
	13	0.24189	0.09115	0.25854	0.12746	0.25691	0.12174
0.02387	26	0.35041	0.30408	0.35247	0.30276	0.34754	0.30041
	52	0.39679	0.40076	0.40692	0.42242	0.39838	0.40460
S&P US 500	4	−0.12126	−0.33386	−0.12387	−0.34077	−0.13386	−0.36154
	13	0.10412	0.14843	0.10293	0.13658	0.10405	0.13861
−0.13630	26	0.32506	0.70820	0.34276	0.75345	0.33539	0.73123
	52	0.31947	0.78032	0.24698	0.62356	0.31258	0.77149
S&P Global 1200	4	0.04634	−0.24378	0.04790	−0.24157	0.04531	−0.24698
	13	0.30869	0.31363	0.31104	0.31682	0.30869	0.31328
−0.02988	26	0.54255	0.83725	0.55551	0.86181	0.55473	0.86418
	52	0.55324	0.99952	0.57734	1.07424	0.56968	1.04838

- using prediction intervals to assess whether, or not, we are achieving portfolios with a constant return

Regression of individual asset returns against time

Notwithstanding that regression against time is necessary in our approach we can explore whether, when individual asset returns are regressed against time, we encounter significant linear regressions.

To provide insight into this question we took each of the assets (stocks) in our largest instance, the S&P Global 1200, and for successive in-sample periods of $h = 6, 13, 26$ weeks over the entire set of asset returns, performed a regression of asset return against time. For each regression we computed the p-value and compared it with 0.05 (indicating whether the regression was significant at the 5% level or not). Table 3.7 shows the results obtained. If there was no dependence between asset returns and time then we would expect (purely by chance) to have 5% of p-values ≤ 0.05.

	$h = 6$	$h = 13$	$h = 26$
Number of p-values	378154	371202	358301
Number of p-values ≤ 0.05	50680	23290	15229
Percentage of p-values ≤ 0.05	13.402	6.274	4.250

Table 3.7: Regression against time, p-value count

It can be seen from Table 3.7 however that for $h = 6, 13$ weeks we have a higher percentage of p-values ≤ 0.05 then we would expect by chance. Given the number of observations (observed p-values) involved, over 350000 observations in each case, these results are highly significant and indicate that, taken collectively, more assets show significant linear dependence on time than we would expect from pure chance. As might be expected the longer the in-sample period (h) the less effect we see. The explanation for this is that in the short-term (for this instance $h = 6, 13$ weeks) a linear regression against time can be used to model asset returns. However over longer time periods such a model is not sustainable. Note here that many of the results given in this chapter deal with short in-sample periods such as $h = 6, 13$ weeks.

Type II errors

Recall that in order to check whether, or not, we have achieved a constant return per time period we above adopted hypothesis testing and in a number of the tables presented conducted hypothesis tests on regression slopes (as to whether regression slope was significantly different from zero or not), at specified significance levels. If we judge that the regression slope is zero then this indicates that (statistically) we have evidence for a constant return per time period.

Formally the null hypothesis used was H_0:regression slope is zero versus the alternative hypothesis H_1:regression slope is different from zero, so a two-sided hypothesis test. Rejecting H_0 when it is true is known as a Type I error and the significance level is the probability of making a Type I error. In hypothesis testing a Type II error relates to accepting (i.e. not rejecting) H_0 when it is false and so we may be interested in the probability of making such an error.

For simplicity we shall restrict our analysis of Type II errors purely to the first table of results presented, Table 3.1. Values for the probability of a Type II error seen in this section were calculated using the R statistical programming language and the pwr package due to Champely (2013), based on Cohen (1988). Table 3.8 gives the results obtained.

To explain Table 3.8 consider the case of ARP-RT and excess return in Table 3.1. Towards the foot of that table we can see that over the 44 cases there were 39 with a slope test success. These are the 39 cases out of the 44 where the hypothesis test H_0:regression slope is zero versus the alternative hypothesis H_1:regression slope is different from zero did not reject H_0 at the significance level of 10% adopted for Table 3.1. As for each of these 39 cases we did not reject H_0 we have the possibility (in each case) that we may have made a Type II error of accepting H_0 when it is false. In other words although we accepted H_0 we should really have rejected it and not recorded a slope test success.

Over the 44 hypothesis tests conducted at the 10% significance level for that column of Table 3.1 we see from Table 3.8 that the average probability of a Type II error was 0.76 (the entry in Table 3.8 corresponding to a significance level of 10%, ARP-RT and excess return). Therefore we have the question as to whether the 39 cases (out of the 44) where we had a slope test success merely reflect Type II errors, or whether there is a genuine effect occurring. To answer this question we calculate the p-value, which here is the probability that we would have seen 39 (or more) cases of slope test success due to Type II errors, each of which occur with probability 0.76 in the 44 hypothesis tests undertaken. This is a simple binomial distribution calculation, namely $\sum_{k=39}^{44}[44!/(k!(44-k)!)](0.76)^k(1-0.76)^{(44-k)}$, which here equates to the 0.030 seen in Table 3.8. The lower the p-value the greater the chance that the effect seen is genuine, i.e. it is not simply a reflection of chance Type II errors occurring.

Table 3.8 gives p-values for varying significance levels, as in Table 3.2. Considering Table 3.8 we can conclude that we are seeing genuine effects (since the p-values are very small) for significance levels of 30% or less.

Significance Level		ARP-RT		ARP-ERT		ARP-RERT	
		return	excess return	return	excess return	return	excess return
10%	# of slope test successes	44	39	44	41	44	41
	Average probability of a Type II error	0.79	0.76	0.80	0.76	0.80	0.76
	p-value	0.000	0.030	0.000	0.003	0.000	0.003
20%	# of slope test successes	40	35	40	35	41	35
	Average probability of a Type II error	0.66	0.63	0.66	0.64	0.66	0.64
	p-value	0.000	0.014	0.000	0.020	0.000	0.020
30%	# of slope test successes	33	31	34	30	34	29
	Average probability of a Type II error	0.55	0.53	0.55	0.53	0.55	0.53
	p-value	0.005	0.014	0.002	0.030	0.002	0.058
50%	# of slope test successes	21	16	20	18	19	18
	Average probability of a Type II error	0.40	0.37	0.41	0.38	0.41	0.38
	p-value	0.186	0.591	0.325	0.400	0.440	0.400

Table 3.8: Probability of Type II errors and p-values

Prediction intervals

An alternative way to judge constancy of return (which does not involve regression) is to use prediction intervals. In-sample we are deciding a portfolio that (ideally) would give the same return in each and every time period. Ideally therefore, out-of-sample, we would find that the decided portfolio has the same (mean) return as it had in-sample.

Since (in the real-world) nothing is ever truly constant we can expand this argument as follows. Suppose we generate a 95% (say) prediction interval (henceforth PI) for the out-of-sample mean portfolio return based on in-sample returns. If, out-of-sample, the mean return that we do achieve lies in this 95% PI then this is evidence that the mean returns, in-sample and out-of-sample, are equivalent.

For readers unfamiliar with the concept a prediction interval, based on in-sample observations, is used to make an inference about either a single future (out-of-sample) observation or a summary statistic relating to out-of-sample observations. It reflects both uncertainty in parameter estimates as well as sampling variation associated with future observations. More technically if we have h in-sample observations of portfolio returns, whose mean is $\overline{r}$ and whose standard deviation is s, then the $100(1-\epsilon)\%$ PI for the out-of-sample mean return based on H out-of-sample return observations is given by $\overline{r} \pm t_{\epsilon/2,h-1}s(\sqrt{1/h+1/H})$. This PI formula, based on the assumption that out-of-sample returns follow the same underlying statistical distribution as in-sample returns, says there is a $100(1-\epsilon)\%$ probability that the mean out-of-sample return will lie in this PI. For simplicity we shall restrict our PI analysis to the results shown in Table 3.1.

Table 3.9 shows the 95% PI analysis for Table 3.1, which used $h = 6$. In Table 3.9

Table 3.9: 95% prediction interval counts

Instance	H	# rebalances	ARP-RT		ARP-ERT		ARP-RERT		
			return	excess return	return	excess return	return	excess return	
S&P Latin America 40	4	98	86	92	85	91	88	94	
	13	31	28	30	28	30	28	30	
	26	16	16	16	16	16	16	16	
	52	8	8	8	8	8	8	8	
S&P Asia 50	4	98	91	93	90	92	90	93	
	13	31	27	24	28	23	28	25	
	26	16	13	12	13	12	13	12	
	52	8	6	5	6	6	6	5	
S&P ASX 50	4	98	88	86	88	86	87	88	
	13	31	26	29	25	27	26	29	
	26	16	16	15	16	15	16	15	
	52	8	8	8	8	6	8	7	
S&P TSX 60	4	98	88	84	89	84	88	82	
	13	31	25	25	25	24	27	24	
	26	16	14	11	15	10	14	10	
	52	8	6	5	6	5	6	5	
S&P UK 125	4	98	86	81	87	82	85	83	
	13	31	26	26	27	26	27	25	
	26	16	14	13	14	13	14	13	
	52	8	7	6	7	5	7	6	
S&P Topix 150	4	98	91	92	91	92	91	90	
	13	31	30	25	30	24	30	27	
	26	16	16	13	16	12	16	11	
	52	8	8	6	8	5	8	6	
S&P Euro Zone 175	4	98	87	84	87	83	87	83	
	13	31	20	23	20	23	20	24	
	26	16	11	11	11	11	11	11	
	52	8	6	5	6	5	6	5	
S&P Euro Plus 225	4	98	87	79	87	80	87	79	
	13	31	21	23	21	23	21	22	
	26	16	11	10	11	9	11	9	
	52	8	6	4	6	4	6	4	
S&P Europe 350	4	98	85	76	86	76	85	77	
	13	31	20	22	21	22	21	22	
	26	16	13	11	13	10	13	10	
	52	8	6	3	6	3	6	3	
S&P US 500	4	98	86	66	86	64	86	66	
	13	31	24	18	24	17	24	18	
	26	16	14	6	14	5	14	7	
	52	8	6	3	6	2	6	3	
S&P Global 1200	4	98	81	68	83	69	83	70	
	13	31	22	17	22	18	22	18	
	26	16	13	6	13	4	14	5	
	52	8	6	3	6	2	6	3	
H = 4		1078	956 (89%)	901 (84%)	959 (89%)	899 (83%)	957 (89%)	905 (84%)	**86%**
H = 13		341	269 (79%)	262 (77%)	271 (79%)	257 (75%)	274 (80%)	264 (77%)	**78%**
H = 26		176	151 (86%)	124 (70%)	152 (86%)	117 (66%)	152 (86%)	119 (68%)	**77%**
H = 52		88	73 (83%)	56 (64%)	73 (83%)	51 (58%)	73 (83%)	55 (63%)	**72%**

we have, for the S&P Latin America 40 with $H = 4$, for example, that the results in Table 3.1 involve 98 rebalances over time. For 86 of these 98 rebalances the actual out-of-sample mean return for ARP-RT lay within the 95% PI as calculated (using the formula shown above) from in-sample data. This therefore is a good indication that out-of-sample returns are following the same underlying statistical distribution as in-sample returns and hence that that the mean return out-of-sample is equal to the mean return in-sample. Note here that since returns can (potentially) be drawn from $(-\infty, +\infty)$ a 95% PI, for example, does not mean that 95% of mean out-of-sample returns should automatically lie in this interval. Summary statistics for the four different H values are given at the foot of Table 3.9. For $H = 4$ and ARP-RT, for example, for 956 out of 1078 rebalances (89% of rebalances) the actual out-of-sample mean return lay within the in-sample derived 95% PI. In general we can see that a high percentage of mean out-of-sample returns lie in the 95% PI derived from in-sample returns.

The actual average values for the mean return ($\overline{r}$) and standard deviation in return (s) as associated with Table 3.9, averaged over all rebalances, are given in Table 3.10. So for example for ARP-RT for the S&P Latin America 40 with $H = 4$ the average (weekly) return (over 98 rebalances) was 0.01289 and the average standard deviation in return was 0.02916. The average (weekly) excess return (over 98 rebalances) was 0.00872 and the average standard deviation in excess return was 0.01928.

One issue with Table 3.9 is that since we are interested in the mean out-of-sample return lying in the 95% PI we should also consider lower prediction levels (such as 90% or below), hence reducing the size of the interval and making it harder for the mean out-of-sample return to lie within the PI. Table 3.11 shows how the results change as we reduce the size of the PI. The 95% PI results in that table are as in Table 3.9, but are repeated there for convenience of comparison. Overall we can see that we still have a high percentage of mean out-of-sample returns lying in the in-sample derived PI (indicating that statistically the mean returns, in-sample and out-of-sample, are equivalent).

In summary here then our results, although produced from an approach based on regression against time, when evaluated from a non-regression (prediction interval) based viewpoint indicate that we are finding portfolios for which in-sample and out-of-sample mean returns are equal. In other words these prediction interval results indicate that we are achieving (mean) constancy of return.

Discussion

In this section we investigated three issues: whether, when asset returns are regressed against time, we encounter significant linear regressions; Type II errors, associated with

Table 3.10: Average mean return and standard deviation in return, averaged over all rebalances

Instance	H	ARP-RT				ARP-ERT				ARP-RERT			
		Ret AV	Ret SD	Exc AV	Exc SD	Ret AV	Ret SD	Exc AV	Exc SD	Ret AV	Ret SD	Exc AV	Exc SD
S&P	4	0.01289	0.02916	0.00872	0.01928	0.01298	0.02952	0.00882	0.01895	0.01291	0.02953	0.00874	0.01941
Latin America 40	13	0.01271	0.03419	0.00892	0.01898	0.01283	0.03425	0.00904	0.01777	0.01241	0.03448	0.00862	0.01850
	26	0.00863	0.03315	0.00674	0.01946	0.00867	0.03399	0.00678	0.01801	0.00933	0.03319	0.00744	0.01893
	52	0.00920	0.03639	0.00495	0.02181	0.01026	0.03707	0.00600	0.01879	0.01114	0.03726	0.00689	0.02064
S&P	4	0.01122	0.02302	0.00892	0.02002	0.01106	0.02326	0.00876	0.02000	0.01103	0.02304	0.00874	0.02002
Asia 50	13	0.01459	0.02592	0.01068	0.01797	0.01442	0.02627	0.01051	0.01794	0.01402	0.02594	0.01011	0.01754
	26	0.01315	0.02415	0.01106	0.01818	0.01332	0.02464	0.01123	0.01800	0.01318	0.02486	0.01109	0.01798
	52	0.01747	0.02602	0.01208	0.01838	0.01799	0.02607	0.01261	0.01789	0.01806	0.02588	0.01268	0.01816
S&P	4	0.00942	0.01968	0.00748	0.01363	0.00934	0.01961	0.00739	0.01352	0.00941	0.01964	0.00747	0.01362
ASX 50	13	0.01127	0.01895	0.00769	0.01403	0.01153	0.01886	0.00795	0.01419	0.01141	0.01899	0.00783	0.01428
	26	0.00866	0.02088	0.00705	0.01360	0.00890	0.02040	0.00730	0.01321	0.00891	0.02043	0.00730	0.01348
	52	0.00953	0.02107	0.00479	0.01241	0.00989	0.02091	0.00515	0.01192	0.00995	0.02075	0.00521	0.01208
S&P	4	0.01353	0.02147	0.01126	0.01823	0.01383	0.02147	0.01156	0.01797	0.01386	0.02143	0.01159	0.01788
TSX 60	13	0.01414	0.02270	0.01149	0.01945	0.01481	0.02312	0.01217	0.01906	0.01439	0.02305	0.01175	0.01910
	26	0.01164	0.02296	0.00902	0.01660	0.01305	0.02344	0.01043	0.01590	0.01299	0.02334	0.01037	0.01638
	52	0.01358	0.02239	0.00925	0.01762	0.01588	0.02306	0.01155	0.01583	0.01544	0.02280	0.01111	0.01764
S&P	4	0.01080	0.01920	0.01013	0.01330	0.01080	0.01928	0.01014	0.01330	0.01081	0.01924	0.01014	0.01324
UK 125	13	0.01075	0.01912	0.01026	0.01505	0.01078	0.01905	0.01029	0.01502	0.01077	0.01904	0.01028	0.01508
	26	0.00774	0.01991	0.01105	0.01600	0.00778	0.01983	0.01109	0.01602	0.00776	0.01989	0.01107	0.01597
	52	0.01056	0.01852	0.01479	0.01617	0.01064	0.01825	0.01487	0.01620	0.01057	0.01856	0.01480	0.01614
S&P	4	0.01041	0.02697	0.00931	0.01638	0.01043	0.02692	0.00933	0.01631	0.01039	0.02698	0.00928	0.01632
Topix 150	13	0.00684	0.02664	0.00894	0.01609	0.00689	0.02669	0.00899	0.01606	0.00682	0.02670	0.00892	0.01614
	26	0.00620	0.02817	0.01083	0.01427	0.00629	0.02830	0.01093	0.01430	0.00626	0.02811	0.01089	0.01427
	52	0.00808	0.02801	0.01258	0.01577	0.00852	0.02792	0.01301	0.01523	0.00848	0.02793	0.01298	0.01518
S&P	4	0.01040	0.01885	0.00943	0.01572	0.01037	0.01896	0.00940	0.01576	0.01037	0.01885	0.00941	0.01576
Euro Zone 175	13	0.01158	0.01987	0.01079	0.01523	0.01148	0.01981	0.01069	0.01521	0.01149	0.01996	0.01070	0.01528
	26	0.00839	0.01745	0.01230	0.01524	0.00831	0.01723	0.01223	0.01525	0.00831	0.01754	0.01223	0.01524
	52	0.00870	0.01793	0.01276	0.01427	0.00860	0.01817	0.01266	0.01427	0.00871	0.01797	0.01277	0.01421
S&P	4	0.01050	0.01896	0.00948	0.01488	0.01050	0.01893	0.00948	0.01486	0.01049	0.01897	0.00948	0.01489
Euro Plus 225	13	0.01195	0.02049	0.01118	0.01481	0.01198	0.02033	0.01121	0.01489	0.01198	0.02056	0.01121	0.01492
	26	0.00881	0.01801	0.01251	0.01507	0.00870	0.01804	0.01240	0.01494	0.00879	0.01813	0.01249	0.01507
	52	0.00883	0.01894	0.01295	0.01413	0.00877	0.01899	0.01289	0.01414	0.00875	0.01913	0.01288	0.01414
S&P	4	0.01108	0.01819	0.01025	0.01300	0.01105	0.01817	0.01021	0.01298	0.01105	0.01824	0.01022	0.01300
Europe 350	13	0.01198	0.01944	0.01135	0.01348	0.01205	0.01936	0.01141	0.01352	0.01199	0.01946	0.01136	0.01351
	26	0.00849	0.01860	0.01205	0.01348	0.00891	0.01847	0.01247	0.01349	0.00887	0.01845	0.01243	0.01353
	52	0.00974	0.01856	0.01390	0.01321	0.00973	0.01872	0.01389	0.01334	0.00973	0.01872	0.01390	0.01325
S&P	4	0.01206	0.01949	0.01206	0.01291	0.01221	0.01961	0.01221	0.01286	0.01220	0.01955	0.01220	0.01285
US 500	13	0.01310	0.02146	0.01254	0.01423	0.01309	0.02166	0.01253	0.01411	0.01331	0.02159	0.01275	0.01428
	26	0.00869	0.02115	0.01122	0.01138	0.00891	0.02139	0.01144	0.01090	0.00858	0.02126	0.01112	0.01106
	52	0.00994	0.01873	0.01198	0.01044	0.01042	0.01903	0.01246	0.01007	0.01022	0.01897	0.01226	0.01019
S&P	4	0.01297	0.01676	0.01248	0.01261	0.01297	0.01677	0.01248	0.01253	0.01291	0.01670	0.01243	0.01256
Global 1200	13	0.01393	0.01858	0.01342	0.01234	0.01392	0.01861	0.01341	0.01229	0.01397	0.01857	0.01346	0.01230
	26	0.01075	0.01788	0.01347	0.01179	0.01100	0.01784	0.01372	0.01166	0.01055	0.01782	0.01327	0.01166
	52	0.01162	0.01703	0.01406	0.01100	0.01176	0.01715	0.01420	0.01070	0.01166	0.01704	0.01410	0.01083
H = 4		0.01139	0.02107	0.00996	0.01545	0.01141	0.02114	0.00998	0.01537	0.01140	0.02111	0.00997	0.01541
H = 13		0.01208	0.02249	0.01066	0.01561	0.01216	0.02255	0.01075	0.01546	0.01205	0.02258	0.01064	0.01554
H = 26		0.00920	0.02203	0.01066	0.01501	0.00944	0.02214	0.01091	0.01470	0.00941	0.02209	0.01088	0.01487
H = 52		0.01066	0.02214	0.01128	0.01502	0.01113	0.02230	0.01175	0.01440	0.01116	0.02227	0.01178	0.01477

falsely accepting the null hypothesis when testing a regression slope for significance; and using prediction intervals to assess whether, or not, we are achieving portfolios with a

Table 3.11: Prediction interval counts

% PI	H	# rebalances	ARP-RT		ARP-ERT		ARP-RERT		
			return	excess return	return	excess return	return	excess return	
95%	4	1078	956 (89%)	901 (84%)	959 (89%)	899 (83%)	957 (89%)	905 (84%)	**86%**
	13	341	269 (79%)	262 (77%)	271 (79%)	257 (75%)	274 (80%)	264 (77%)	**78%**
	26	176	151 (86%)	124 (70%)	152 (86%)	117 (66%)	152 (86%)	119 (68%)	**77%**
	52	88	73 (83%)	56 (64%)	73 (83%)	51 (58%)	73 (83%)	55 (63%)	**72%**
90%	4	1078	870 (81%)	811 (75%)	873 (81%)	806 (75%)	879 (82%)	807 (75%)	**78%**
	13	341	244 (72%)	215 (63%)	241 (71%)	219 (64%)	243 (71%)	221 (65%)	**68%**
	26	176	139 (79%)	96 (55%)	135 (77%)	90 (51%)	138 (78%)	93 (53%)	**65%**
	52	88	68 (77%)	43 (49%)	66 (75%)	41 (47%)	66 (75%)	41 (47%)	**62%**
75%	4	1078	664 (62%)	607 (56%)	664 (62%)	602 (56%)	667 (62%)	613 (57%)	**59%**
	13	341	190 (56%)	143 (42%)	187 (55%)	142 (42%)	191 (56%)	144 (42%)	**49%**
	26	176	111 (63%)	58 (33%)	108 (61%)	58 (33%)	110 (63%)	58 (33%)	**48%**
	52	88	47 (53%)	35 (40%)	46 (52%)	32 (36%)	46 (52%)	34 (39%)	**45%**
50%	4	1078	382 (35%)	391 (36%)	386 (36%)	368 (34%)	385 (36%)	380 (35%)	**35%**
	13	341	99 (29%)	88 (26%)	94 (28%)	88 (26%)	95 (28%)	91 (27%)	**27%**
	26	176	64 (36%)	40 (23%)	61 (35%)	33 (19%)	61 (35%)	38 (22%)	**28%**
	52	88	29 (33%)	25 (28%)	28 (32%)	21 (24%)	27 (31%)	23 (26%)	**29%**

constant return.

Overall we found evidence that (for short in-sample periods) regressing asset returns against time can be justified for more assets than one would expect by chance. With respect to Type II errors we found clear evidence that our results are not due to an accumulation of Type II errors. With respect to prediction intervals, which are derived from a non-regression standpoint, we found evidence that our results indicate that we are achieving portfolios with the same mean return in-sample as out-of-sample.

3.4.6 Further insight

In this section we investigate a number of components of our approach in an attempt to provide further insight. Specifically we consider: intercept maximisation, the number of assets (K) in the portfolio, and the equally-weighted ($1/N$) portfolio. All of these are considered below.

Intercept maximisation

In the second stage of the three-stage objective presented above we maximise the regression intercept (equation 3.16). This is done subject to a constraint (equation 3.17) upon the absolute value of the regression slope achieved at the first stage. The question therefore arises as to whether this stage actually adds value, i.e. whether given the constraints that are applied there is actually flexibility to significantly increase the regression intercept above the value that it took at the first stage.

Rebalance period	ARP-RT			ARP-ERT			ARP-RERT		
	Minimum intercept second stage	Actual intercept first stage	Maximum intercept second stage	Minimum intercept second stage	Actual intercept first stage	Maximum intercept second stage	Minimum intercept second stage	Actual intercept first stage	Maximum intercept second stage
6	-0.0208531	-0.0026045	0.0130541	-0.0137571	0.0045750	0.0212491	-0.0178608	-0.0025197	0.0171434
58	-0.0349128	-0.0107490	0.0146454	-0.0349267	-0.0107526	0.0144553	-0.0346306	-0.0107508	0.0144322
110	-0.0156035	0.0002821	0.0181195	-0.0153263	0.0002836	0.0170755	-0.0153019	0.0010662	0.0180907
162	-0.0242924	-0.0052888	0.0061846	-0.0243803	-0.0052886	0.0073373	-0.0243129	-0.0037053	0.0067570
214	-0.0046683	-0.0025656	0.0003830	-0.0046615	-0.0025572	0.0002837	-0.0046442	-0.0025656	0.0007464
266	-0.0004425	0.0065422	0.0179883	-0.0021896	0.0065542	0.0187162	-0.0041764	0.0075600	0.0187396
318	-0.0071153	-0.0011524	0.0075390	-0.0065627	-0.0009106	0.0075240	-0.0067069	-0.0011523	0.0074886
370	-0.0051645	0.0068485	0.0195998	-0.0051045	0.0068487	0.0184260	-0.0051627	0.0068486	0.0196511

Table 3.12: Intercept values, S&P Global 1200, $H = 52$

To give insight into this question we present Table 3.12. Here, for simplicity, we focus purely on the largest instance, the S&P Global 1200, associated with an in-sample period of $h = 6$ weeks and out-of-sample periods of $H = 4, 13, 26, 52$ weeks. Table 3.12, which corresponds to the $H = 52$ entry for the S&P Global 1200 in Table 3.1, shows for ARP-RT, ARP-ERT and ARP-RERT the actual intercept value achieved at the first stage and the maximum intercept value achieved at the second stage at each rebalance. For interest we also show the minimum intercept value that is achievable at the second stage. The actual intercept value achieved at the first stage must lie somewhere between these minimum and maximum values, albeit the value given by the optimisation at the first stage will be an arbitrarily decided value somewhere between these two limits. So in Table 3.12 for ARP-RT we can have that at the first rebalance in week 6 the actual intercept at the first stage is -0.0026045. The minimum and maximum values that the intercept can achieve at the second stage are -0.0208531 and 0.0130541 respectively.

In order to decide whether the second stage maximisation is adding value we take the difference between the maximum intercept value and the actual intercept value. This, for Table 3.12 and the ARP-RT, will give eight difference values (of which the first will be $0.0130541 - (-0.0026045) = 0.0156586$) and we can then apply the hypothesis test H_0:differences are equal to zero versus H_1:differences are greater than zero, so a one-sided hypothesis test. When this is done for ARP-RT in Table 3.12 the result is highly significant (a p-value of approximately 0.0004), indicating that the maximisation at the second stage is giving significantly increased intercept values. Conducting the same hypothesis test for ARP-ERT and ARP-RERT in Table 3.12 also indicates that the second stage is giving significantly increased intercept values (p-values also approximately 0.0004). For $H = 4, 13, 26$ for this instance the results (though not shown here) are even more significant. Over these three values of H and all three of ARP-RT, ARP-ERT and ARP-RERT the maximum p-value associated with this hypothesis test was 10^{-7}. Overall then, from this analysis, we can conclude that the second stage objective of intercept maximisation in our three-stage objective is adding value.

Number of assets (K) in the portfolio

In the results presented above we have used a value for K, the number of assets in the portfolio, equal to $0.8N$. In this section we illustrate why we choose this particular value of K and what happens as we change K. For simplicity we concentrate in this section on the results shown in Table 3.1 for the largest instance, the S&P Global 1200, for the ARP-RT associated with an in-sample period of $h = 6$ weeks and varying out-of-sample periods, namely $H = 4, 13, 26, 52$.

Figure 3.3 shows for this instance the average out-of-sample return and the associated risk (standard deviation in return) for ten different values of K ($K = 1.0N, 0.9N, \ldots, 0.2N$, $0.1N$) for the four values of H considered.

It is evident from Figure 3.3 that for $H = 13, 26, 52$ as we decrease K from $1.0N$ we increase the return achieved, but also increase the risk taken. For $H = 4$ the situation is slightly different from the other values of H in that as K decreases from $1.0N$ we initially reduce return, but after $0.5N$ we increase return. However even for this case as K decreases we increase risk.

One clear point to emerge from Figure 3.3, which deals with the largest instance with ten different values for K and four different values for H (so 40 out-of-sample cases), is that the three-stage model we have used in this chapter is capable of producing portfolios with good out-of-sample return performance across a wide range of different portfolio sizes.

Given the behaviour seen in Figure 3.3 there is no one unique value of K that can be recommended. For this instance, for all values of H, decreasing K from $1.0N$ to $0.1N$ will increase risk, but for most values of H decreasing K from $1.0N$ to $0.1N$ will also increase return. Hence there is a tradeoff to be made, how much increased risk is an investor prepared to accept for increased return?

Given this implicit tradeoff we, in the results presented above, adopted a value for K of $0.8N$. ***This means that the results given above are presented in a conservative light*** (since Figure 3.3 indicates that higher levels of return are available for smaller values of K), but also means that we are implicitly adopting a conservative attitude to risk (in

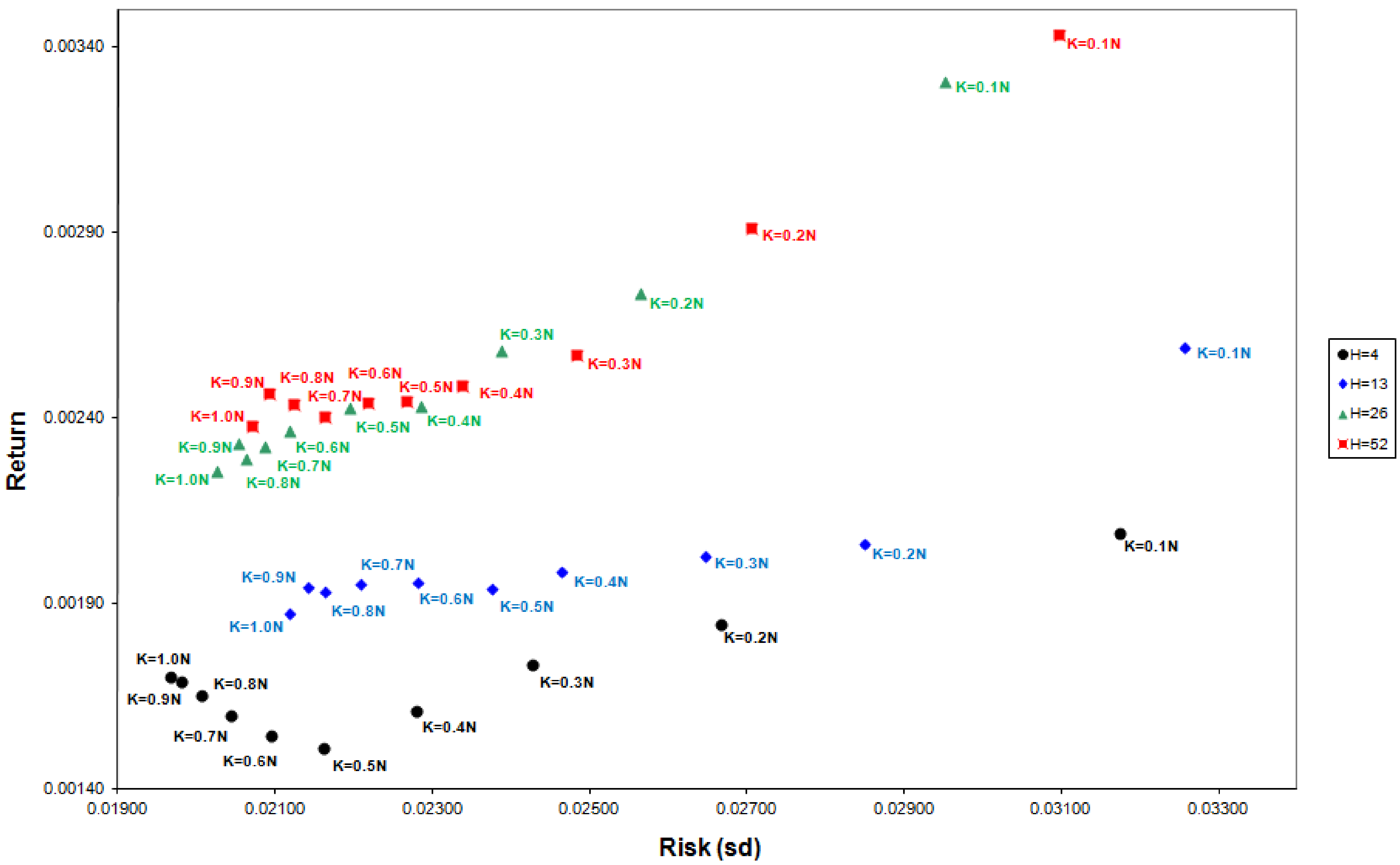

Figure 3.3: Varying K for the S&P Global 1200

that we do not want values of risk which are too high).

Clearly readers of this chapter might make other choices for K. The point we would emphasise here is that, for simplicity, we have adopted a consistent value for K above rather than attempt to present tables showing results for many different values of K. We would make the point here though that our approach is capable of allowing an investor to quickly investigate (since computational times are low, of the order of seconds) any values of K that may be of interest to them.

Equally-weighted ($1/N$) portfolio

With regard to portfolio optimisation in a Markowitz mean-variance framework some authors (e.g. De Miguel et al. (2009)) have argued for use of a non-optimised portfolio based on allocating the same investment proportion to each of the N assets. This portfolio is often referred to as the equally-weighted, or $1/N$, portfolio. The question therefore arises whether such a portfolio will perform better than the (optimised) portfolio produced by the approach given in this chapter.

To give insight into this question we present Figure 3.4. Here, for simplicity, we focus purely on the returns from ARP-RT, the first column of returns in Table 3.1. For each of the 44 cases associated with that column we computed the average out-of-sample return and standard deviation in return for the $1/N$ portfolio. Note here that as Table 3.1 involves rebalancing over time we also rebalanced our $1/N$ portfolio over time.

Figure 3.4 has on the vertical axis average return from our approach (as in the first return column of Table 3.1) minus the average return from the $1/N$ portfolio. On the horizontal axis in Figure 3.4 we have the standard deviation in return from our approach minus the standard deviation in return from the $1/N$ portfolio. With these measures one quadrant in Figure 3.4, namely the upper-left quadrant containing 11 cases, corresponds to cases where our approach produces a result that dominates the $1/N$ approach (since in that quadrant the return difference is positive, so our approach has a better return, and the risk difference is negative, so our approach has a lower risk). The lower-right quadrant, containing 5 cases, corresponds to cases where our approach produces a result that is dominated by the $1/N$ approach. The upper-right and lower-left quadrants in Figure 3.4, containing in total $(44 - 11 - 5) = 28$ cases, correspond to quadrants where there is a risk-return tradeoff to be made between the portfolios given by our approach and the $1/N$ portfolio. In the upper-right quadrant for example our approach gives a better return, but at a higher risk, than the $1/N$ portfolio.

Although Figure 3.4 shows that in 11 of the 44 cases our approach dominates the $1/N$ approach there are 28 cases where no clear conclusion can be drawn, since the approach

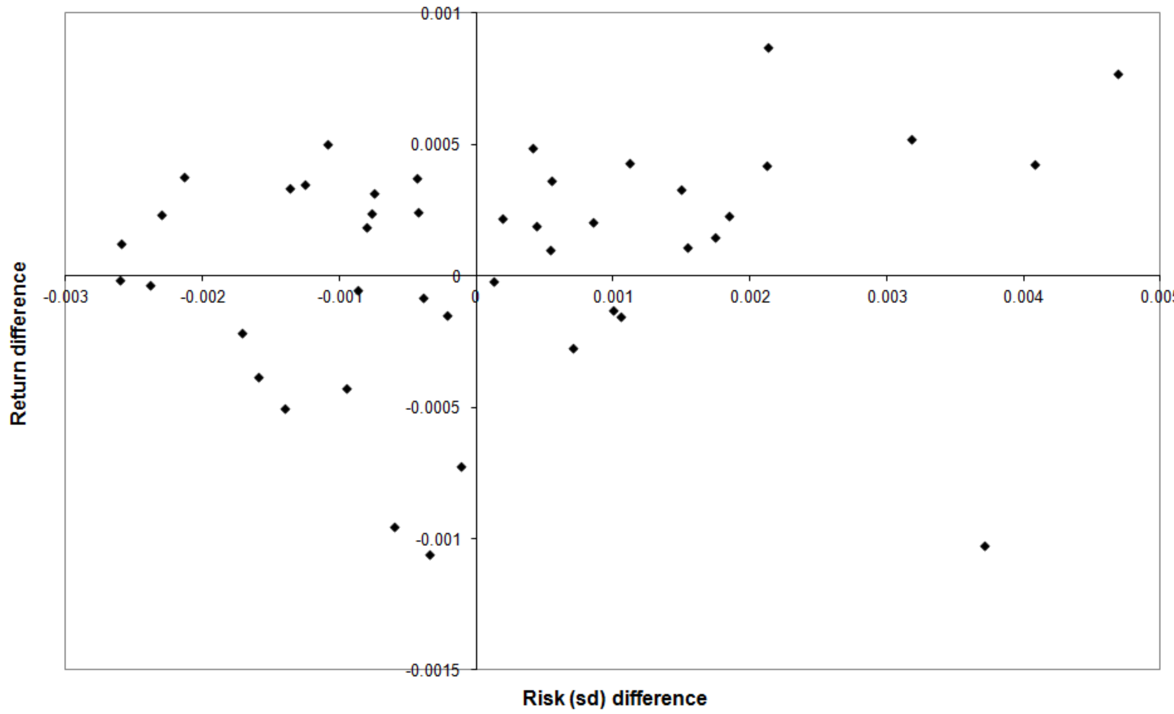

Figure 3.4: ARP-RT, our approach compared with the $1/N$ portfolio

preferred depends upon deciding a risk-return tradeoff. In order to gain insight into such a risk-return tradeoff we present Sharpe ratios for the $1/N$ portfolio in Table 3.13. Sharpe ratios, involving as they do the risk-free rate, are one way to perform a risk-return tradeoff without the investor making a personal value judgement as to the relative worth of return as against risk. The higher the Sharpe ratio, the better.

The Sharpe ratios in Table 3.13 are directly comparable with the Sharpe ratios seen in Table 3.4. So for the S&P Latin America 40 with $H = 4$, with regard to the series of returns obtained, the Sharpe ratio from our approach for ARP-RT in Table 3.4 is 0.54814, whereas the Sharpe ratio for the $1/N$ approach in Table 3.13 is 0.57463, so slightly better. With regard to the series of excess returns obtained, the Sharpe ratio from our approach for ARP-RT in Table 3.4 is -0.31723 whereas the Sharpe ratio for the $1/N$ approach in Table 3.13 is -0.43000, so slightly worse.

Table 3.14 gives the number of cases, out of 44, where the $1/N$ portfolio approach has a better (higher) Sharpe ratio than our approach as in Table 3.4, and vice-versa, where our approach is better. For the return column for ARP-RT in Table 3.14 there are 16 Sharpe ratios in Table 3.13 that are better than those in Table 3.4, and by implication 44-16=28 cases in Table 3.4 where the Sharpe ratio is better than that in Table 3.13. Overall we can see that for each of the six columns in Table 3.14 our approach out-performs the

Table 3.13: Sharpe ratios, $1/N$ portfolio

Instance	H	return	excess return
S&P Latin America 40	4	0.57463	−0.43000
	13	0.56399	−0.48071
0.57367	26	0.54736	−0.53535
	52	0.53730	−0.59960
S&P Asia 50	4	0.73190	0.26619
	13	0.70504	0.21540
0.40393	26	0.70950	0.22341
	52	0.72837	0.27523
S&P ASX 50	4	0.86711	0.65426
	13	0.86174	0.62166
0.36924	26	0.86918	0.63802
	52	0.89287	0.70219
S&P TSX 60	4	0.84937	0.19442
	13	0.85743	0.20541
0.42601	26	0.90225	0.26370
	52	1.01176	0.37219
S&P UK 125	4	0.23813	0.25302
	13	0.22612	0.22177
−0.05901	26	0.28677	0.34373
	52	0.35368	0.47310
S&P Topix 150	4	0.37386	0.38686
	13	0.36257	0.36101
0.09071	26	0.38663	0.44828
	52	0.47576	0.72370
S&P Euro Zone 175	4	0.24576	0.06452
	13	0.23941	0.04298
0.04990	26	0.27653	0.13583
	52	0.33359	0.25246
S&P Euro Plus 225	4	0.26375	0.08314
	13	0.25621	0.05522
0.06440	26	0.29277	0.15408
	52	0.34214	0.26295
S&P Europe 350	4	0.27056	0.20272
	13	0.25964	0.16502
0.02387	26	0.30555	0.29388
	52	0.36381	0.43305
S&P US 500	4	0.28592	0.61640
	13	0.24427	0.52750
−0.13630	26	0.27190	0.58455
	52	0.30763	0.64860
S&P Global 1200	4	0.47589	0.81066
	13	0.45016	0.74231
−0.02988	26	0.48780	0.82378
	52	0.56958	0.96974

$1/N$ approach. In addition the summary values indicate that in 160 out of the 264 cases considered (so 61% of cases) our approach out-performs the $1/N$ approach.

Since these results are for the case $\varepsilon_i = 0.25/K, \delta_i = 2/K$ where the maximum investment in any particular asset is restricted to be at most $2/K$ we also show in Table 3.14 results for the case $\varepsilon_i = 0.25/K, \delta_i = 0.10$. In this case 25% of the total investment must (as a minimum) be spread amongst the K assets, but the remaining 75% is free, and indeed up to 10% of the total investment could be placed in one (or more) individual assets. We can see that for this case for each of the six columns in Table 3.14 our approach again out-performs the $1/N$ approach. The summary values given indicate that in 157 out of the 264 cases considered (so 59% of cases) our approach out-performs the $1/N$ approach.

Table 3.14: Sharpe ratios comparison

	ARP-RT		ARP-ERT		ARP-RERT		
	return	excess return	return	excess return	return	excess return	
$\varepsilon_i = 0.25/K, \delta_i = 2/K$							
Number of cases where $1/N$ is better	16	17	16	18	19	18	**104**
Number of cases where our approach is better	28	27	28	26	25	26	**160**
$\varepsilon_i = 0.25/K, \delta_i = 0.10$							
Number of cases where $1/N$ is better	17	19	16	19	17	19	**107**
Number of cases where our approach is better	27	25	28	25	27	25	**157**

3.4.7 Discussion

In this chapter we have presented computational results considering eleven different problem instances with three parameters that the investor can choose: the number of assets (K) in the portfolio, the length of the in-sample period (h) from which data is drawn to decide the portfolio using our three-stage model and the length of the out-of-sample period (H) for which the portfolio is held before rebalancing occurs.

For simplicity we limited the number of different combinations of K:h:H that we examined in detail for eleven problem instances. However we hope that we have presented sufficient results above to convince the reader that:

- our three-stage model is of benefit. To address this point we would mention the low computation time and positive out-of-sample results and Sharpe ratios produced. In addition we would stress here that within our three-stage model we can address finding either an absolute return portfolio (ARP-RT), or a relative return portfolio (ARP-ERT) or a mixed portfolio (ARP-RERT).

- the individual stages in our three-stage model are each of value. To address this point we would mention the logic of minimising the absolute value of the regression slope at the first stage; the results showing that maximising the intercept at the second stage leads to significantly better intercepts; the logic of including a transaction cost minimisation third stage.

- the results obtained are not simply an artefact of the portfolio size (K) adopted. To address this issue we would mention the results presented showing that over a wide range of K values our three-stage model is capable of producing portfolios with good out-of-sample return performance.

- the results obtained are better than those from an alternative approach. To address this issue we would mention the results presented showing that over all eleven instances our three-stage model produces portfolios with better out-of-sample Sharpe ratios than an alternative approach based on an equally-weighted ($1/N$) portfolio.

3.4.8 Further research

It is clear that the work presented in this chapter can be taken further. In particular there are three future lines of inquiry that we would highlight here: higher-frequency price data, a rebalance/liquidate/reinvest trading strategy and comparison with other models previously presented in the literature.

With respect to higher-frequency price data note that we have used weekly asset price data. The data set we used in this chapter was previously used in Meade & Beasley (2011) and had the advantage that it had been manually adjusted to account for changes in index composition, removing susceptibility to the influence of survivor bias.

Whilst obtaining price data (say at daily frequency) from commercial databases such as Datastream is relatively easy there is significant effort involved in adjusting such data to account for index composition changes. These typically need to be included since in deciding a portfolio (of any type) the asset universe needs to be defined. Frequently asset universes are related to index composition, as in the work presented here, where the universe at any point in time was composed of the assets that were in the considered index at that point in time. In other words in historical back-testing, such as presented in this chapter, one needs to know the index composition at any point in time.

Therefore, due to the effort required to assemble a survivor bias free set of daily price data for sufficient markets/indices to establish results with any confidence, we believe that investigating the performance of our three-stage model using such data remains a topic for future research.

A rebalance/liquidate/reinvest trading strategy is a strategy where, at each decision point, an investor either: rebalances to a new portfolio (such as might be produced by any of the models we have given above); or liquidates the entire existing portfolio for cash; or reinvests from cash into a portfolio. For example an investor already holding a portfolio might liquidate the portfolio if they believe that asset returns in the immediate future will be negative. An investor holding cash from a previously liquidated portfolio might choose to reinvest if they believe that asset returns in the immediate future will be high enough. This strategy could be evaluated by considering cash as an asset with zero return.

One way to incorporate such a trading strategy into the work we have presented here is to look at the prediction interval for the mean out-of-sample return at each decision point. Taking the example of possible liquidation, if at a specified prediction level this prediction interval does not lie above zero (or an appropriate risk-free rate) then the investor liquidates the portfolio, since the prediction is that out-of-sample returns will not be high enough.

We do however believe that an investigation as to whether (or not) such a trading strategy would (after accounting for transaction cost) generate higher returns (or produce better Sharpe ratios) than the pure rebalance only strategy we have used in this chapter is a matter for future research.

With respect to comparison with other models previously presented in the literature then, as already mentioned above, different papers deal with different models, designed for different purposes, and it is difficult to compare them one to another.

At the core of the three-stage model we have presented is an absolute return focus in that (ideally) we would find a portfolio that achieves a constant return per time period. Other models in the literature typically have a different focus. For example they may focus on: simple Markowitz mean-variance based portfolio optimisation; or Markowitz mean-variance based portfolio optimisation but with cardinality constraints; or simple index tracking; or index tracking but with additional constraints upon tracking error such as those based upon conditional value at risk; or enhanced indexation.

It is true that one can regard all of these models simply as black boxes which at any given point in time produce a portfolio from in-sample data that one can evaluate in terms of out-of-sample performance. The model that best rewards the investor in terms of out-of-sample performance, based on historical back-testing, is the one that the (practical) investor might choose.

However, given the number of different models that exist, we believe that such a comparison remains a topic for future research.

3.5 Conclusions

In this chapter we have considered the problem of selecting an absolute return portfolio. We presented a three-stage mixed-integer zero-one program for the problem that explicitly considers transaction costs associated with trading. We extended our approach to present models for enhanced indexation (relative return) portfolios and for portfolios that are a mix of absolute and relative return. Computational results were given for portfolios derived from universes defined by S&P international equity indices which indicated that all three models produced good quality results and that the computation time required was not significant.

Chapter 4

Market neutral portfolios

4.1 Introduction

A market neutral portfolio is a portfolio of financial assets that (ideally) exhibits performance independent from that of an underlying market as represented by a benchmark index. Funds that claim to adopt market neutral strategies are common. Traditionally such funds hold both long and short positions. A long position in a particular asset is where a fund purchases a number of units of that asset (e.g. stocks, shares). A fund will typically hold a long position if it takes the view that the price of the asset will rise. If their view is correct then they will see a gain in the value of the asset (which can, via sale, be turned into a cash profit).

A short position is where the fund finds a current holder of a particular asset, borrows it from them and sells the asset immediately in the market. They do this because they think that the price of the asset will fall, enabling them to buy the asset back later at a lower price and return it to the original lender. If their view that the price will fall is correct this would yield a profit in cash terms. Holding a short position is also known as shorting, or short selling.

In this chapter we consider the problem of constructing a market neutral portfolio where we can hold both long and short positions in assets. We formulate this problem as a mixed-integer nonlinear program, minimising the absolute value of the correlation between portfolio return and index return.

The contribution of this chapter is:

- to present a model for market neutral portfolios that directly addresses, via mixed-integer nonlinear programming, the minimisation of correlation between the portfolio chosen and the return on the benchmark index
- to present computational results, for test instances involving up to 1200 assets, which

indicate that our mixed-integer nonlinear approach is computationally feasible

- to show, for the test problems considered, that the results for the model proposed out-perform an alternative approach based on minimising the absolute value of regression slope (the zero-beta approach)

To the best of our knowledge this chapter is the first in the literature to directly address correlation minimisation.

The structure of this chapter is as follows. In Section 4.2 we give our formulation of the problem of constructing a market neutral portfolio. In Section 4.3 we give computational results for constructing market neutral portfolios for eleven different problem instances derived from universes defined by S&P international equity indices. In Section 4.4 we present relevant discussions regarding our model. Finally in Section 4.5 we present our conclusions.

4.2 Problem formulation

4.2.1 Overview

Market neutral portfolios seek to avoid market risk. One way (as in Patton (2009)) to define a market neutral portfolio (henceforth MNP) is as a portfolio that has zero correlation with the market (for example, as represented by a benchmark index).

In the model presented in this chapter we adopt the view that in seeking a MNP we are looking for a portfolio that ideally has zero correlation with a market benchmark. In practice we may not find such a portfolio, but if we could find a portfolio whose returns have zero correlation with market returns, we would then have an ideal MNP.

In this chapter we formulate a mixed-integer nonlinear program for the problem of selecting a MNP. Nonlinearities exist in the model since our objective is to find a correlation of zero between our portfolio and the market. Nonlinear formulations are often computationally challenging; in order to find (locally) optimal solutions we used the Minotaur solver (Leyffer et al. (2013)), a toolkit for solving mixed-integer nonlinear optimisation problems. In the following sections we give our notation and present the constraints and objective that we used to find a MNP. Our model is a general one in that it allows both long and short positions.

4.2.2 Notation

We observe over time $0, 1, 2, \ldots, T$ the value of N assets. We are interested in selecting the best set of K assets to hold (where $K \leq N$), as well as their appropriate quantities.

Let:

V_{it}	be the value (price) of one unit of asset i at time t
r_{it}	be the single period continuous time return for asset i at time t, i.e. $r_{it} = \ln(V_{it}/V_{it-1})$
I_t	be the value of the benchmark market index at time t
R_t	be the single period continuous time return for the index at time t, i.e. $R_t = \ln(I_t/I_{t-1})$
$\bar{R}$	be the average return on the index, i.e. $\bar{R} = \sum_{t=1}^{T} R_t/T$
C	be the total cash available (≥ 0) at time T to invest in the MNP
C^L, C^S	be the limits (> 0) on the total invested in long/short positions at time T
ε_i^L, ε_i^S	be the lower limits ($0 \leq \varepsilon_i^L, \varepsilon_i^S \leq 1$) on the proportion of C^L and C^S respectively invested in long/short positions in asset i if any position is taken
δ_i^L, δ_i^S	be the upper limits ($0 \leq \delta_i^L, \delta_i^S \leq 1$) on the proportion of C^L and C^S respectively invested in long/short positions in asset i

Then our decision variables are:

x_i^L, x_i^S	the number of units (≥ 0) of asset i that we choose to hold in long/short positions respectively
w_i^L, w_i^S	the "proportion" of the initial investment (cash) C held in long/short positions in asset i
z_i^L, z_i^S	$\begin{cases} 1 & \text{if any of asset } i \text{ is held in long/short positions in the MNP} \\ 0 & \text{otherwise} \end{cases}$

Without significant loss of generality (since the sums of money involved are large) we allow $[x_i^L]$, $[x_i^S]$ to take fractional values.

Observe that "proportion", relating to w_i^L and w_i^S, is in inverted comma's. When we have shorting (as opposed to a long-only portfolio) the usual long-only proportion interpretation changes. When we have a long-only portfolio we are using our initial investment C to purchase assets. As such the proportion invested in any asset must always be a fraction (≤ 1) of C.

However when we allow shorting we can increase the amount that we have available for purchases. Here we first borrow from an intermediary x_i^S units of asset i ($i = 1, \ldots, N$), sell them immediately in the market for a price of V_{iT} each and hence have a cash sum $C + \sum_{i=1}^{N} x_i^S V_{iT}$ to invest in long positions. For this reason the amount invested in any asset as a proportion of C can exceed 1, unlike the usual long-only interpretation of a proportion as being ≤ 1.

4.2.3 Constraints

The constraints associated with our formulation of the problem of finding a MNP are:

$$w_i^L = x_i^L V_{iT}/C, \quad i = 1, \ldots, N \tag{4.1}$$

$$w_i^S = x_i^S V_{iT}/C, \quad i = 1, \ldots, N \tag{4.2}$$

$$\varepsilon_i^L z_i^L \leq x_i^L V_{iT}/C^L \leq \delta_i^L z_i^L, \quad i = 1, \ldots, N \tag{4.3}$$

$$\varepsilon_i^S z_i^S \leq x_i^S V_{iT}/C^S \leq \delta_i^S z_i^S, \quad i = 1, \ldots, N \tag{4.4}$$

$$z_i^L + z_i^S \leq 1, \quad i = 1, \ldots, N \tag{4.5}$$

$$\sum_{i=1}^{N} (z_i^L + z_i^S) = K \tag{4.6}$$

$$\sum_{i=1}^{N} x_i^L V_{iT} \leq C^L \tag{4.7}$$

$$\sum_{i=1}^{N} x_i^S V_{iT} \leq C^S \tag{4.8}$$

$$\sum_{i=1}^{N} w_i^L - \sum_{i=1}^{N} w_i^S = 1 \tag{4.9}$$

Equations (4.1) and (4.2) define the "proportion" w_i^L and w_i^S of the MNP invested in asset i. These variables (although strictly unnecessary as they can be substituted out algebraically) are introduced here to ease the mathematics presented. Equations (4.3) and (4.4) ensure that if an asset i is not held in a long/short position in the MNP ($z_i^L = 0, z_i^S = 0$) then x_i^L, x_i^S are also zero. If asset i is held then these equations ensure that the proportions of C^L, C^S held respect the limits defined. Equation (4.5) says that we cannot hold a long and short position in the same asset simultaneously. Equation (4.6) ensures that there are exactly K assets in the MNP. Equations (4.7) and (4.8) ensure that we respect the limits on the total invested in long/short positions.

Equation (4.9) is the monetary balance equation and is equivalent to:

$$\sum_{i=1}^{N} x_i^L V_{iT} = C + \sum_{i=1}^{N} x_i^S V_{iT} \tag{4.10}$$

where we assume $C \neq 0$. This equation says that at time T we first generate cash $\sum_{i=1}^{N} x_i^S V_{iT}$ by shorting, add that to C and then purchase long positions costing $\sum_{i=1}^{N} x_i^L V_{iT}$.

Note here that amending the constraints we have given above so as to have long-only portfolios without shorting is trivial (simply remove the variables associated with shorting).

4.2.4 Objective function

According to our definition of a MNP, we seek zero correlation between portfolio return and market benchmark return. Statistically, correlation is defined by a nonlinear formula and it ranges from -1 to $+1$. Define the following additional decision variables:

C_t	the value of the MNP at time $t = 0, \ldots, T$
p_t	the MNP log-return at time $t = 1, \ldots, T$
$\bar{p}$	the average return on the MNP

These variables are given by:

$$C_t = \sum_{i=1}^{N} x_i^L V_{it} - \sum_{i=1}^{N} x_i^S V_{it}, \quad t = 0, \ldots, T \tag{4.11}$$

$$p_t = \ln(C_t / C_{t-1}), \quad t = 1, \ldots, T \tag{4.12}$$

$$\bar{p} = \sum_{t=1}^{T} p_t / T \tag{4.13}$$

In Equation (4.11) we have that at time t the MNP has $x_i^L,\ i = 1, \ldots, N$ in long positions, collectively worth $\sum_{i=1}^{N} x_i^L V_{it}$. The short positions $x_i^S,\ i = 1, \ldots, N$ represent obligations which have to be repaid (since in shorting, short selling, we borrow assets and have to return them). Collectively the short positions represent a (monetary) repayment of $\sum_{i=1}^{N} x_i^S V_{it}$ and so the value of the MNP at time t is as given in Equation (4.11). Equation (4.12) defines the log-returns for the MNP and Equation (4.13) the average return.

It is clear from Equation (4.11) that the presence of short selling may mean that the value of the MNP becomes negative (or zero). In either case this would mean that the corresponding return (Equation (4.12)) is not defined. Numerically this would result in the optimisation software we are using giving an error. For this reason we impose the constraint that $C_t,\ t = 0, \ldots, T$ is greater than (or equal to) some small positive value. In the computational results reported later we used $C = 1000000$ and imposed the constraint that $C_t \geq 10,\ t = 0, \ldots, T$.

In our model to find a MNP our objective is to minimise the absolute value of the correlation between MNP return and benchmark index return. Utilising the formula for correlation, the Pearson product-moment coefficient (which can be found in any statistics textbook), our objective is:

$$\text{minimise} \left| \frac{\sum_{t=1}^{T}(p_t - \bar{p})(R_t - \bar{R})}{\sqrt{\sum_{t=1}^{T}(p_t - \bar{p})^2 \sum_{t=1}^{T}(R_t - \bar{R})^2}} \right| \tag{4.14}$$

This minimisation is subject to the constraints given in Equations (4.1)-(4.9), (4.11)-(4.13). Our model for a MNP is a mixed-integer nonlinear program which is non-convex (due to Equation (4.12)). Note here that although the objective (Equation (4.14)) can be simplified, since the term $\sum_{t=1}^{T}(R_t - \bar{R})^2$ is a constant and so can be dropped from the objective without affecting the optimal solution, we retain it in the form given for consistency.

Note also that we can replace this nonlinear modulus objective by a linear objective (if so desired) by introducing, in a standard way, a single additional variable ($E \geq 0$) and two nonlinear constraints. The model then becomes:

$$\text{minimise } E \tag{4.15}$$

subject to (4.1)-(4.9), (4.11)-(4.13) and:

$$E \geq \frac{\sum_{t=1}^{T}(p_t - \bar{p})(R_t - \bar{R})}{\sqrt{\sum_{t=1}^{T}(p_t - \bar{p})^2 \sum_{t=1}^{T}(R_t - \bar{R})^2}} \tag{4.16}$$

$$E \geq -\frac{\sum_{t=1}^{T}(p_t - \bar{p})(R_t - \bar{R})}{\sqrt{\sum_{t=1}^{T}(p_t - \bar{p})^2 \sum_{t=1}^{T}(R_t - \bar{R})^2}} \tag{4.17}$$

4.2.5 Other constraints

The model we have given above for the problem of deciding a market neutral portfolio can be easily extended to deal with additional (user specified) constraints that may be encountered in practice. Examples of such constraints that are especially relevant to MNPs are given below.

Dollar neutral

The term "dollar neutral" with respect to a MNP is used to signify that, over a set Q of assets at time T, the value of the long positions is equal to the value of the short positions.

The constraint that enforces this is:

$$\sum_{i \in Q} x_i^L V_{iT} = \sum_{i \in Q} x_i^S V_{iT} \tag{4.18}$$

For example, we might impose dollar neutrality on a set of assets within one or more economic sectors (e.g. banking, telecommunications) when forming an equity/stock based MNP. Note here however that attempting to impose dollar neutrality on the entire MNP, i.e. using $Q = [1, \ldots, N]$, would from Equation (4.10) explicitly require $C = 0$. This, depending upon how portfolio returns are calculated, can introduce complexities into the model, since potentially we start with $C = 0$ and generate via judicious long/short positions some money from nothing. For a discussion as to this issue see Khandani & Lo (2007).

It is a simple matter to extend the concept of dollar neutrality to ensure that the net position over a set Q of assets, i.e. $\sum_{i \in Q} x_i^L V_{iT} - \sum_{i \in Q} x_i^S V_{iT}$, lies within prescribed limits.

Long/short fix

In a long/short fix MNP we specify, as a proportion of C, the amount that can be held in long/short positions. Letting α (where $0 \leq \alpha \leq 1$) be the proportion of C that can be held in short positions we have:

$$C^S = \alpha C \tag{4.19}$$

$$C^L = (1 + \alpha)C \tag{4.20}$$

If $\alpha = 0.30$, for example, we would hold a maximum of 30% of C in short positions and a maximum of 130% of C in long positions. Note here that if we wish to exactly specify the amount held in long/short positions all that needs to be done is to change the inequalities in Equations (4.7) and (4.8) to equalities.

There are many funds that limit their exposure with a proportion $\alpha = 0.30$ and they are usually referred to as 130/30 funds (see Lo & Patel (2008); Tol & Wanningen (2009, 2011)). Other values for α are also encountered, e.g. $\alpha = 0.20$ for 120/20 funds (Jacobs & Levy (2007)) or $\alpha = 0.50$ for 150/50 funds.

Regulation T

Regulation T is a rule that applies in the USA to investors of certain types who make use of shorting. It essentially prevents the proceeds from shorting being used for long purchases. In a rule of this type the entire proceeds from shorting (so $\sum_{i=1}^{N} x_i^S V_{iT}$), plus an additional fraction γ of this amount, must be set aside. This additional fraction is to cover the risk that the price of assets that have been shorted will rise. For example, regulation T uses $\gamma = 0.5$. If a rule of this type applies then Equations (4.10) and (4.9) become:

$$\sum_{i=1}^{N} x_i^L V_{iT} = C - \gamma \sum_{i=1}^{N} x_i^S V_{iT} \tag{4.21}$$

$$\sum_{i=1}^{N} w_i^L + \gamma \sum_{i=1}^{N} w_i^S = 1 \tag{4.22}$$

So here in Equation (4.21) the initial cash sum C available for long purchases is reduced by the fraction needed to support shorting (c.f. Equation (4.10) where the initial cash sum C is supplemented by the proceeds from shorting). Equation (4.11), defining the value of the portfolio at time t, changes to:

$$C_t = \sum_{i=1}^{N} x_i^L V_{it} - \sum_{i=1}^{N} x_i^S V_{it} + (1+\gamma) \sum_{i=1}^{N} x_i^S V_{iT}, \quad t = 0, \ldots, T \tag{4.23}$$

In this equation the value of the portfolio at time t is composed of a long position ($\sum_{i=1}^{N} x_i^L V_{it}$); a short position which is an obligation to be repaid ($\sum_{i=1}^{N} x_i^S V_{it}$); plus the proceeds from shorting that have been set aside ($\sum_{i=1}^{N} x_i^S V_{iT}$); plus an additional fixed amount, taken at time T from C, of $\gamma \sum_{i=1}^{N} x_i^S V_{iT}$.

In-sample returns

The model presented above seeks to find a MNP with a correlation of zero between portfolio and market returns, but does not seek to achieve high returns in-sample (over the period $t = 1, \ldots, T$). In order to improve in-sample MNP return (which may, or may not, improve out-of-sample return) we can add:

$$\bar{p} \geq \bar{R} \tag{4.24}$$

This constraint ensures that the average return from the portfolio chosen will be at least that of the benchmark index.

4.2.6 Zero-beta approach

One alternative approach to constructing a market neutral portfolio is the zero-beta approach. Here beta relates to the slope of the regression line when portfolio returns are regressed against the returns from the benchmark index. In this approach portfolio return is defined using the usual linear approximation as a weighted sum of asset returns:

$$p_t = \sum_{i=1}^{N} (w_i^L - w_i^S) r_{it}, \quad t = 1, \ldots, T \tag{4.25}$$

Let $\hat{\beta}_i$ be the regression slope when the returns r_{it} from asset i are regressed against index returns R_t (in-sample, so over the period $t = 1, \ldots, T$). The zero-beta model involves minimising the absolute value of the (in-sample) portfolio regression slope when portfolio returns p_t are regressed against index returns R_t. From Equation (4.25) this is:

$$\text{minimise} \ \mid \sum_{i=1}^{N} (w_i^L - w_i^S) \hat{\beta}_i \mid \tag{4.26}$$

This minimisation is subject to the constraints given in Equations (4.1)-(4.9). Although the modulus in Equation (4.26) makes the objective nonlinear it can be linearised in the same fashion as for Equation (4.14) above, meaning that the zero-beta model is a mixed-integer linear program that computationally can be easily solved.

We would note in passing here that some authors in the literature (see Alexander (1977); Baele & Londono (2013); Black (1972)) consider the minimum variance zero-beta portfolio, where the objective function is to minimise the variance in portfolio return and a constraint is imposed to ensure that beta is precisely zero. The difficulty with this approach is that, whilst it is possible to find such a portfolio when the problem is effectively unrestricted, as additional restrictions (such as considered above, Equations (4.3)-(4.8)) are added it can become impossible to make beta precisely zero. Such approaches typically also use the linear approximation (Equation (4.25)) referred to above.

As discussed above our approach directly minimises |correlation|, as given in Equation (4.14). Suppose we were to regress the actual returns (p_t, Equation (4.12)) from the portfolio chosen against index returns R_t, so without any approximation (such as Equation (4.25)). The regression slope β that we would get is related to correlation by correlation $= \beta[\text{sd}(R_t)/\text{sd}(p_t)]$. It is clear from this that minimising |correlation| is equivalent to minimising $|\beta/\text{sd}(p_t)|$, since $\text{sd}(R_t)$ is a constant. Our approach therefore is related to minimisation of the regression slope β obtained without approximation, but is not exactly the same, since we include the term $\text{sd}(p_t)$ involving the standard deviation in MNP return (again without approximation).

In summary here therefore the key differences between the zero-beta model and our MNP approach are:

- the zero-beta approach adopts a linear approximation of portfolio return (Equation (4.25)) in order to achieve a computationally efficient mixed-integer linear model

- we adopt a nonlinear model that involves no approximation of portfolio return and directly addresses the minimisation of correlation

4.3 Computational results

In this section we present computational results for our approach to constructing market neutral portfolios. We used an Intel Xeon CPU E5-2640 @ 2.50GHz with 64GB of RAM with Linux as the operating system. The code was written in C++ with Minotaur 0.1.1 (Leyffer et al. (2013)) used as the mixed-integer nonlinear solver for our MNP model: optimise Equation (4.15) subject to Equations (4.1)-(4.9), (4.11)-(4.13), (4.16), (4.17), (4.19), (4.20) and, where appropriate below, Equation (4.24).

4.3.1 Minotaur

Minotaur is a recently developed toolkit for solving mixed-integer nonlinear programs. It has two main solvers, one based on nonlinear branch-and-bound denoted as bnb, the other an implementation of a QP-diving algorithm. In the work presented in this chapter we used the bnb solver.

With reference to our choice of the Minotaur solver the reader may be aware that there are many mixed-integer nonlinear solvers available (e.g. 25 different solvers are listed in Burer & Letchford (2012); Bussieck & Vigerske (2011)). Our choice of Minotaur was guided by results given recently by Mittelmann (2012). He compared a number of well-known solvers (such as BARON, Couenne, KNITRO, LINDO-global, Minotaur and SCIP) on hundreds of test problems (including those in Bussieck et al. (2003); MINLP Library (2013)). Of the solvers he considered he recommended Minotaur as a first choice, followed by KNITRO and then SCIP. Of these solvers two (Minotaur and SCIP) are free to download and use permanently for academic purposes, KNITRO requires a licence. We have investigated both Minotaur and SCIP. Overall SCIP, for the specific test problems we considered, did not out-perform Minotaur and so in the results presented below we use Minotaur.

Clearly a different solver may give better results than Minotaur for the test problems we considered, but in our view the results presented below are of sufficient quality (both in terms of computation time and in terms of solution optimality) to make further exploration of different solvers a secondary issue.

Independently of the chosen solver, mixed-integer nonlinear programs are computationally hard to solve and given the current state of software packages benefit can be gained by the user manually amending a formulation. Here an issue that caused instabilities with regard to Minotaur was the inclusion of the square root term associated with correlation. We rewrote Equations (4.16) and (4.17) by adding variable $b \geq \tau$ (where τ was a small positive constant, we used 10^{-8}). These variables were defined using:

$$b^2 - \left(\sum_{t=1}^{T} (p_t - \bar{p})^2 \sum_{t=1}^{T} (R_t - \bar{R})^2 \right) = 0 \tag{4.27}$$

$$E \times b \geq \left(\sum_{t=1}^{T} (p_t - \bar{p})(R_t - \bar{R}) \right) \tag{4.28}$$

$$E \times b \geq - \left(\sum_{t=1}^{T} (p_t - \bar{p})(R_t - \bar{R}) \right) \tag{4.29}$$

4.3.2 Data and methodology

In our computational experimentation we adopted successive periodic rebalancing over time and we used the same real-world historical weekly data as described in Section 3.4.1, the reader may refer to that section for details on the methodology.

With regard to parameter values we, unless otherwise stated, set the in-sample period $h = 13$ and the out-of-sample period $H = 13$. We set $C = 1000000$ corresponding to an initial investment of US\$1 million. We examined $\alpha = [0, 0.30, 0.50]$, which represent Long Only, 130/30 and 150/50 portfolios. Unless otherwise stated we used $K = N$ as the number of assets in the MNP; $\varepsilon_i^L = \varepsilon_i^S = 0$ and $\delta_i^L = \delta_i^S = 1$ for $i = 1, \ldots, N$ as the proportion limits for each asset in the MNP. With these values the solver has complete freedom to choose the best (lowest correlation) in-sample MNP amongst all possible solutions.

We imposed a time limit of $\max[2N, 400]$ seconds for each rebalance of our MNP, if the time limit is reached before Minotaur terminates we retrieve the best feasible solution found so far. In order to speed up Minotaur we provided a feasible solution (when available from the previous rebalance) as a warm start.

4.3.3 Results, in-sample

Table 4.1 compares the three variations investigated (Long Only, 130/30 and 150/50) for each test instance. The ***correlation*** columns give the average in-sample correlation (Equation (4.14)) value over all rebalances. Given the total number of time intervals (400 weeks) and our choice of $H = 13$, there were a total of $\lfloor 400/H \rfloor = 30$ rebalances. The ***return*** columns are the average in-sample weekly continuous time returns and the ***excess*** columns represent the average in-sample excess over the average index weekly continuous time return. The ***t(s)*** columns are the average times (in seconds) taken by Minotaur at each rebalance. The first of the bottom rows of Table 4.1 shows the average value for each column, whilst the last row shows the annualised returns and excess returns (in percentage per annum). The conversion of average weekly returns to yearly returns is given by the expression $100(\exp(\text{return} \times 52) - 1)$. For instance, the average in-sample return for Long Only is 0.00386 which corresponds to a yearly return of $100(\exp(0.00386 \times 52) - 1) = 22.23\%$.

Table 4.1: Summary of in-sample results

Instance	Long Only				130/30				150/50			
	correlation	return	excess	t(s)	correlation	return	excess	t(s)	correlation	return	excess	t(s)
S&P Latin America 40	0.59099	0.00511	0.00201	0.06	0.01537	0.00988	0.00679	2.73	−0.01183	0.00933	0.00623	1.91
S&P Asia 50	0.17830	0.00165	−0.00058	0.03	0.00000	0.00225	0.00002	4.58	0.05679	0.00730	0.00507	4.12
S&P ASX 50	0.28930	0.00019	−0.00169	0.03	0.00001	0.00273	0.00084	4.04	0.04547	0.01084	0.00896	3.12
S&P TSX 60	0.07486	0.00268	0.00053	0.04	0.00002	0.01670	0.01455	3.75	−0.00006	0.01185	0.00970	3.02
S&P UK 125	0.00198	0.00202	0.00177	0.07	0.01214	0.00684	0.00658	8.91	−0.00001	0.00557	0.00531	6.11
S&P Topix 150	0.02336	0.00409	0.00317	0.07	0.11975	0.02368	0.02277	40.71	0.10280	0.02804	0.02712	30.41
S&P Euro-Zone 175	0.03409	0.00470	0.00436	0.10	0.00001	0.00714	0.00680	27.77	0.00000	0.00775	0.00741	41.50
S&P Euro-Plus 225	0.00349	0.00097	0.00058	0.12	−0.00001	0.00539	0.00501	54.37	−0.00002	0.00555	0.00516	72.40
S&P Europe 350	0.02503	0.00576	0.00542	0.16	0.00911	0.00940	0.00906	115.79	0.00001	0.00560	0.00526	95.84
S&P US 500	0.00018	0.00542	0.00537	0.21	0.11566	0.04635	0.04630	166.75	0.04556	0.03467	0.03463	312.39
S&P Global 1200	0.13051	0.00988	0.00954	0.52	−0.00001	0.00927	0.00893	890.12	0.00000	0.04695	0.04662	969.73
Average	**0.12292**	**0.00386**	**0.00277**	**0.13**	**0.02473**	**0.01269**	**0.01160**	**119.96**	**0.02170**	**0.01577**	**0.01468**	**140.05**
Yearly Return (%)		**22.23**	**15.49**			**93.46**	**82.80**			**127.06**	**114.55**	

In terms of execution time, Minotaur is able to solve the Long Only case very efficiently. The addition of short selling increases execution time substantially, but at manageable levels. Considering all rebalances for all instances (so $30 \times 11 = 330$ cases), the time limit was reached in only three cases: once for the S&P US 500 with 150/50 and twice for the S&P Global 1200, once with 130/30 and once with 150/50.

The consequence of adding short selling can also be seen in Table 4.1 where the average in-sample correlation for Long Only is in most cases considerably higher than the average in-sample correlation for 130/30 and 150/50. The introduction of short selling, adding flexibility to the MNP that can be constructed, improves in-sample correlation (as we would expect).

Even though in the results seen in Table 4.1 no special modification (such as Equation (4.24)) was made to improve returns, the in-sample MNPs found consistently outperformed the index. For Long Only the average weekly excess return was 0.00277 (15.49% yearly) and for 9 of the 11 instances the Long Only MNPs outperformed the index. The introduction of short selling improved the in-sample returns even further, where the MNPs for all instances outperformed their respective indices and the average weekly excess returns were 0.01160 and 0.01468 (82.80% and 114.55% yearly) respectively for 130/30 and

150/50.

Minotaur may not be able to guarantee solution optimality for non-convex problems, but given our formulation we know that a MNP which has zero in-sample correlation with the index must be an optimal solution. Table 4.2 shows for each instance how many times (out of the 30 rebalances) a correlation of zero was found (within a tolerance of 10^{-4}). In most cases Minotaur was able to find an optimal solution to the problem. Moreover, we cannot tell if a correlation different from zero means a suboptimal solution. For Long Only the average number of proven optima was 20.3, whilst for 130/30 and 150/50 the average number of proven optima was approximately 26.

Table 4.2: Number of optimal solutions found

Instance	Long Only	130/30	150/50
S&P Latin America 40	4	27	28
S&P Asia 50	16	28	25
S&P ASX 50	7	29	25
S&P TSX 60	22	27	27
S&P UK 125	27	27	27
S&P Topix 150	26	24	23
S&P Euro-Zone 175	22	27	27
S&P Euro-Plus 225	27	28	28
S&P Europe 350	26	24	28
S&P US 500	25	19	22
S&P Global 1200	21	27	26
Average	**20.3/30**	**26.1/30**	**26.0/30**

4.3.4 Results, out-of-sample

In this section we analyse how our MNPs perform out-of-sample. Table 4.3 presents the out-of-sample results. It has the same format as Table 4.1, except for the absence of the time column.

Compared to the in-sample results we would expect deterioration in out-of-sample correlation, and this can indeed be seen when comparing Table 4.3 with Table 4.1. The average out-of-sample correlation was 0.52379 for Long Only, 0.40825 for 130/30 and 0.40467 for 150/50. Again we can see that the inclusion of short selling improves the correlation. In terms of returns, even though there was also deterioration as compared to in-sample, out-of-sample the MNPs still performed better than the index in most instances: namely in 8, 7 and 7 of the 11 instances for Long Only, 130/30 and 150/50 respectively. Whilst the in-sample returns were higher when we included short selling,

Table 4.3: Summary of out-of-sample results

Instance	Long Only			130/30			150/50		
	correlation	return	excess	correlation	return	excess	correlation	return	excess
S&P Latin America 40	0.66847	0.00539	0.00236	0.46693	0.00221	−0.00082	0.41558	0.00225	−0.00078
S&P Asia 50	0.48222	0.00086	−0.00124	0.44551	0.00083	−0.00127	0.40066	0.00278	0.00068
S&P ASX 50	0.50134	0.00045	−0.00135	0.46967	−0.00100	−0.00281	0.44516	0.00152	−0.00028
S&P TSX 60	0.35245	0.00068	−0.00156	0.23426	0.00458	0.00234	0.34455	0.00312	0.00089
S&P UK 125	0.57860	0.00244	0.00210	0.48193	0.00379	0.00346	0.55159	0.00275	0.00241
S&P Topix 150	0.42840	0.00122	0.00056	0.40607	0.00447	0.00381	0.37390	−0.00002	−0.00068
S&P Euro-Zone 175	0.57333	0.00443	0.00362	0.43123	0.00262	0.00181	0.44034	0.00443	0.00361
S&P Euro-Plus 225	0.58486	0.00246	0.00158	0.42279	0.00357	0.00269	0.37137	0.00251	0.00163
S&P Europe 350	0.55033	0.00390	0.00323	0.44945	0.00291	0.00224	0.49490	0.00161	0.00094
S&P US 500	0.55907	0.00308	0.00312	0.40525	0.00528	0.00531	0.37808	0.00318	0.00322
S&P Global 1200	0.48257	0.00632	0.00596	0.27766	−0.00168	−0.00205	0.23523	−0.00141	−0.00177
Average	**0.52379**	**0.00284**	**0.00167**	**0.40825**	**0.00251**	**0.00134**	**0.40467**	**0.00207**	**0.00090**
Yearly Return (%)		**15.91**	**9.07**		**13.94**	**7.22**		**11.36**	**4.79**

Long Only had on average a better out-of-sample performance with an average return of 0.00284 (15.91% yearly), compared to 0.00251 and 0.00207 (13.94% and 11.36% yearly) for 130/30 and 150/50.

4.3.5 Comparison with the zero-beta portfolio

Table 4.4 presents results comparing the model we have presented in this chapter with the zero-beta model: optimise Equation (4.26) subject to Equations (4.1)-(4.9),(4.19),(4.20). To ensure a consistent comparison these results are just for the first optimisation (so no rebalancing is performed, since if we were to consider rebalancing our model and the zero-beta model might well pursue different trajectories through time due to choosing different portfolios). The three columns associated with our model show the in-sample correlation and in-sample regression slope (β) associated with the portfolio chosen, and the time taken to generate that portfolio. Note that the correlation value (ignoring the sign) corresponds to the optimised objective function (Equation (4.14)) value.

For the zero-beta model the first column shows the optimised objective function (Equation (4.26)) value and the fourth column the time taken. In all but one instance (Long Only, S&P ASX 50) the objective function value is effectively zero. For the portfolio chosen by the zero-beta model we show in the second column its in-sample correlation and in the third column its in-sample regression slope. Note that the difference between the first and third columns here is due to the fact that in the zero-beta model a linear approximation has been applied through the use of Equation (4.25). This typically means that the actual regression slope, as computed directly from the portfolio chosen (third column), will differ from the optimised objective function value (first column). Comparing

Table 4.4: Comparison with the zero-beta model

Case	Instance	Our model			Zero-beta model			
		Correlation	Slope β	t(s)	Objective value	Correlation	Slope β	t(s)
Long Only	S&P Latin America 40	−0.00001	−0.00001	0.02	0.00000	−0.01698	−0.01888	0.00
	S&P Asia 50	−0.00001	−0.00001	0.02	0.00000	−0.00534	−0.00336	0.00
	S&P ASX 50	0.02282	0.03173	0.02	0.03170	0.02282	0.03173	0.00
	S&P TSX 60	−0.00001	−0.00001	0.02	0.00000	0.06855	0.03400	0.00
	S&P UK 125	−0.00000	−0.00001	0.03	0.00000	−0.00323	−0.01068	0.00
	S&P Topix 150	−0.00000	−0.00000	0.04	0.00000	0.18121	0.18166	0.00
	S&P Euro Zone 175	−0.00002	−0.00006	0.03	0.00000	0.00010	0.00028	0.00
	S&P Euro Plus 225	0.00001	0.00002	0.04	0.00000	0.01169	0.01499	0.00
	S&P Europe 350	−0.00002	−0.00004	0.06	0.00000	0.00325	0.00606	0.01
	S&P US 500	0.00002	0.00006	0.08	0.00000	−0.00221	−0.00586	0.01
	S&P Global 1200	0.00001	0.00002	0.16	0.00000	−0.00205	−0.00322	0.03
	Mean Absolute Value	**0.00208**	**0.00291**	**0.05**	**0.00288**	**0.02886**	**0.02825**	**0.00**
130/30	S&P Latin America 40	−0.00001	−0.00001	1.06	0.00000	−0.03983	−0.02680	0.01
	S&P Asia 50	0.14055	0.12169	0.47	0.00000	0.14160	0.12278	0.01
	S&P ASX 50	0.00001	0.00001	1.80	0.00000	0.00558	0.01120	0.01
	S&P TSX 60	0.00002	0.00001	0.95	0.00000	0.01151	0.00890	0.01
	S&P UK 125	0.00000	0.00001	4.37	0.00000	0.00183	0.00529	0.03
	S&P Topix 150	0.00000	0.00000	6.90	0.00000	0.15222	0.22825	0.03
	S&P Euro Zone 175	−0.00001	−0.00004	4.79	0.00000	0.00170	0.00470	0.03
	S&P Euro Plus 225	−0.00000	−0.00002	21.32	0.00000	0.02318	0.08521	0.04
	S&P Europe 350	−0.00001	−0.00004	20.55	0.00000	0.00237	0.00886	0.08
	S&P US 500	−0.00002	−0.00003	298.64	0.00000	−0.00976	−0.01271	0.15
	S&P Global 1200	−0.00001	−0.00006	863.86	0.00000	0.00158	0.00759	0.49
	Mean Absolute Value	**0.01279**	**0.01108**	**111.34**	**0.00000**	**0.03556**	**0.04748**	**0.08**
150/50	S&P Latin America 40	−0.00002	−0.00001	0.69	0.00000	0.04291	0.02867	0.01
	S&P Asia 50	0.00001	0.00000	0.93	0.00000	−0.02135	−0.01287	0.01
	S&P ASX 50	0.00001	0.00002	2.15	0.00000	−0.03718	−0.06362	0.01
	S&P TSX 60	−0.00001	−0.00001	1.14	0.00000	−0.01568	−0.01805	0.01
	S&P UK 125	0.00000	0.00000	5.27	0.00000	0.01252	0.05414	0.03
	S&P Topix 150	−0.00000	−0.00000	6.80	0.00000	−0.02562	−0.05244	0.03
	S&P Euro Zone 175	0.00002	0.00005	9.28	0.00000	−0.00151	−0.00446	0.04
	S&P Euro Plus 225	0.00000	0.00002	8.92	0.00000	0.02283	0.12897	0.04
	S&P Europe 350	−0.00002	−0.00005	55.54	0.00000	−0.03009	−0.10259	0.09
	S&P US 500	0.00001	0.00002	133.03	0.00000	0.01183	0.03039	0.15
	S&P Global 1200	−0.00000	−0.00002	772.58	0.00000	0.04193	0.24992	0.49
	Mean Absolute Value	**0.00001**	**0.00002**	**90.58**	**0.00000**	**0.02395**	**0.06783**	**0.08**

our model and the zero-beta model then:

- with regard to correlation in all 33 instances |correlation| is lower (or equal) for our model than for the zero-beta model
- with regard to regression slope (β) we can see by comparing the second column under our model with the actual regression slope associated with the zero-beta model (third column for that model) that in all 33 instances the portfolios chosen by our model have lower (or equal) $|\beta|$ values than the portfolios chosen by the zero-beta model

Clearly (especially for the larger instances) our model requires more time than the zero-beta model. However, given its superiority (for the instances considered) with regard to the actual in-sample correlations and regression slopes associated with the portfolios chosen, we can reasonably claim that the portfolios produced by our model out-perform those produced by the zero-beta model.

4.3.6 Comparison with market neutral S&P 500 funds

In this section we compare the performance of our approach to selecting MNPs with funds that define themselves as market neutral.

We searched for market neutral funds which had one of our S&P indices (Figure 3.1), as a benchmark. Since our data spans 7 years from 1999 to 2006, market neutral funds were required to have at least 50 prices during this time period. Data was collected from Thomson Reuters Datastream (2013), where we found a total of 7 funds whose benchmark was the S&P 500 index and which met our requirements. Market neutral funds that use the other S&P indices as benchmark and which were active prior to 2006 are relatively rare and thus we only present results for market neutral S&P 500 funds.

A comparison between our MNP approach and these S&P 500 funds (labelled Funds 1 to 7) is shown in Table 4.5. Once again the ***correlation***, ***return*** and ***excess*** columns for our MNPs are out-of-sample averages. The column labelled # is the number of weeks of data we have for each fund. All funds have a starting date somewhere within the period 1999 to 2006. To achieve a fair comparison, in each row of Table 4.5 the correlation, return and excess for our approach only includes the time period over which the respective fund was active.

Table 4.5: Comparison with market neutral S&P 500 funds

Fund	#	Fund performance			Long Only			130/30			150/50		
		correlation	return	excess	correlation	return	excess	correlation	return	excess	correlation	return	excess
Fund 1	168	0.34564	0.00052	-0.00099	0.64355	0.00424	0.00273	0.29689	0.00454	0.00303	0.37325	0.00343	0.00192
Fund 2	168	0.33655	0.00042	-0.00110	0.64355	0.00424	0.00273	0.29689	0.00454	0.00303	0.37325	0.00343	0.00192
Fund 3	168	0.34434	0.00042	-0.00110	0.64355	0.00424	0.00273	0.29689	0.00454	0.00303	0.37325	0.00343	0.00192
Fund 4	234	0.14031	-0.00025	-0.00071	0.56119	0.00402	0.00355	0.38520	0.00218	0.00172	0.38975	0.00282	0.00235
Fund 5	234	0.13141	-0.00030	-0.00077	0.56119	0.00402	0.00355	0.38520	0.00218	0.00172	0.38975	0.00282	0.00235
Fund 6	324	0.08832	-0.00002	0.00039	0.56335	0.00319	0.00360	0.38670	-0.00102	-0.00061	0.32419	0.00101	0.00142
Fund 7	285	0.23301	-0.00018	-0.00055	0.57507	0.00385	0.00348	0.39191	-0.00021	-0.00058	0.35985	0.00205	0.00168
Average		**0.23137**	**0.00009**	**-0.00069**	**0.59878**	**0.00397**	**0.00320**	**0.34853**	**0.00239**	**0.00162**	**0.36904**	**0.00271**	**0.00194**
Yearly Return %			**0.47**	**-3.52**		**22.93**	**18.04**		**13.29**	**8.79**		**15.13**	**10.61**

Overall these funds achieve better out-of-sample correlations than our approaches. However for funds 1, 2 and 3 both our 130/30 and 150/50 MNPs achieve competitive correlations (slightly better for 130/30, slightly worse for 150/50).

Of course a practical investor in a market neutral fund will be interested in the return gained and here the comparison between these 7 funds and our MNPs is marked. Only one fund (fund 6) generated return in excess of the S&P 500 benchmark over the relatively long-term timescale considered (from 168 to 324 weeks, 3.2 to 6.2 years). Even there that excess return (0.00039 weekly, 2.05% yearly) was not large.

By contrast our MNPs were able to outperform the S&P 500 index in the vast majority of the (different) time periods considered. In only 2 of the 21 MNP cases in Table 4.5 was underperformance against the benchmark seen. Long Only had on average the best out-of-sample performance with an average weekly excess return of 0.00320 (18.04% yearly), compared to 0.00162 and 0.00194 (8.79% and 10.61% yearly) for 130/30 and 150/50.

Of course our approach to constructing MNPs does not include transaction costs. The funds shown in Table 4.5 would have incurred transaction costs associated with trading (but these are not known, nor is the frequency at which these funds trade known).

Various authors have given different estimates of transaction cost for long-only purchases (see Meade & Beasley (2011) for a discussion as to this). Given that the results in Table 4.5 are for a holding period H of 13 weeks then rebalancing only occurs 4 times per year. As an approximation therefore unless the transaction costs associated with rebalancing a Long Only portfolio exceed $18.04/4 = 4.5\%$ our Long Only MNP would (on average) have out-performed the index. Estimates given in Meade & Beasley (2011) indicate that such a high transaction cost is unlikely, particularly since here we are considering assets in the S&P 500 which are very liquid and frequently traded. Overall therefore it seems reasonable to conclude that our Long Only MNP would have out-performed the market and the funds considered.

With respect to the cost of shorting Cohen et al. (2007) provide a discussion. A similar approximation as to transaction cost can be done for 130/30 and 150/50 as was done above for Long Only. However, in our view, the differences in estimated costs of shorting given in the literature make any conclusions that we might draw somewhat speculative (particularly as our holding period of $H = 13$ weeks implies that any assets we borrow and short will be absent from their lender for a considerable period of time).

4.3.7 Variations

Above we have presented computational results for eleven different problem instances with a number of parameters that the user can set. Our MNP approach is very flexible and has a significant number of parameters which can be varied. For example, K the number of assets in the MNP and H the length of the out-of-sample period.

For simplicity we did not investigate all possible variations. However Table 4.6 presents in-sample summary statistics, averaged over the eleven problem instances, for a number of variations. The first row in Table 4.6 is as the summary row at the foot of Table 4.1 and is repeated here for ease of comparison. The next row presents the result when we include Equation (4.24) which ensures that in-sample return is at least index return. The next two rows are as the first two rows, except here for an out-of-sample holding period of $H = 4$ weeks. In the final three rows the number of assets K in the MNP is $0.2N$. In the last but one row we apply Equation (4.24), whilst in the final row we apply proportions ($\varepsilon_i^L = \varepsilon_i^S = 0.25/K$; $\delta_i^L = \delta_i^S = 2/K$). Table 4.7 is as Table 4.6 except that it relates to out-of-sample summary statistics. In Table 4.7 the first row is as the summary row at the foot of Table 4.3.

Table 4.6: In-sample summary statistics, different variations

K	H	Extra	Long Only				130/30				150/50			
			correlation	return	excess	t(s)	correlation	return	excess	t(s)	correlation	return	excess	t(s)
N	13	-	0.12292	0.00386	0.00277	0.13	0.02473	0.01269	0.01160	119.96	0.02170	0.01577	0.01468	140.05
N	13	$\bar{p} \geq \bar{R}$	0.12187	0.01260	0.01151	0.07	0.02363	0.05700	0.05593	89.18	0.01367	0.06836	0.06727	93.51
N	4	-	0.10841	0.00357	0.00244	0.06	0.00509	0.01067	0.00953	134.15	0.01332	0.01201	0.01088	142.26
N	4	$\bar{p} \geq \bar{R}$	0.19567	0.01371	0.01258	0.06	0.01123	0.03908	0.03794	130.66	0.01395	0.05633	0.05522	122.87
0.2N	13	-	0.11368	0.00428	0.00319	44.09	0.02061	0.01129	0.01021	155.89	0.00526	0.01484	0.01376	153.75
0.2N	13	$\bar{p} \geq \bar{R}$	0.10166	0.01682	0.01576	21.99	0.01988	0.08469	0.08368	110.11	0.00936	0.14861	0.14755	92.93
0.2N	13	Proportions	0.11266	0.00195	0.00086	50.07	0.00290	0.00382	0.00286	165.27	0.00017	0.00511	0.00408	159.42

Table 4.7: Out-of-sample summary statistics, different variations

K	H	Extra	Long Only			130/30			150/50		
			correlation	return	excess	correlation	return	excess	correlation	return	excess
N	13	-	0.52379	0.00284	0.00167	0.40825	0.00251	0.00134	0.40467	0.00207	0.00090
N	13	$\bar{p} \geq \bar{R}$	0.51053	0.00228	0.00111	0.40074	0.00206	0.00089	0.34118	0.00023	-0.00093
N	4	-	0.51837	0.00282	0.00165	0.44827	0.00201	0.00084	0.39624	0.00119	0.00002
N	4	$\bar{p} \geq \bar{R}$	0.47898	0.00313	0.00196	0.39580	0.00036	-0.00080	0.33535	0.00111	-0.00005
0.2N	13	-	0.53371	0.00315	0.00198	0.39303	0.00239	0.00122	0.38484	0.00065	-0.00052
0.2N	13	$\bar{p} \geq \bar{R}$	0.45883	0.00349	0.00232	0.35415	0.00224	0.00107	0.26937	0.00140	0.00023
0.2N	13	Proportions	0.72186	0.00247	0.00131	0.63721	0.00199	0.00082	0.60861	0.00206	0.00089

Examining Tables 4.6 and 4.7 there is a mixed picture. For example, comparing the nine cases where we include Equation (4.24), with the equivalent cases where we do not, shows that for all of these cases Equation (4.24) results in better returns in-sample, but in only three cases do we see better returns out-of-sample. Using $H = 4$ instead of $H = 13$, so a shorter holding period out-of-sample, does not seem to result in any improvement in returns, either in-sample or out-of-sample. With respect to the use of proportions the results indicate a significant increase in out-of-sample correlation.

Using $K = 0.2N$ (without proportions) instead of $K = N$, so having fewer assets in the MNP, is better in four out of six cases both in-sample and out-of-sample. So here there is some evidence, based on averaging over eleven instances in each of the six cases, that restricting the number of assets in the MNP can be beneficial.

4.4 Discussions

4.4.1 Choice of regression for the ARP model

In Chapter 3 we presented the ARP model, which searches for a portfolio with a constant return per time period. In that model, we used a regression based strategy as opposed to a correlation based one.

In fact, we could not have chosen correlation as a suitable nonlinear objective function for the ARP model, due to the nature of the constant return assumption. For instance, the correlation between time, the benchmark we chose for the ARP (e.g. time represented as numbers from 1 to h), and a hypothetical investment whose returns are perfectly constant would be undefined. As the variance of the constant return investment is zero, we cannot calculate the Pearson correlation coefficient.

Now suppose our returns are as close as possible to being perfectly constant, i.e. $h-1$ constant returns and a single return that is different from the others by a very small amount. The correlation of this example does not follow a predictable pattern, and this can be verified by artificially constructing simple examples.

4.4.2 Pearson correlation coefficient

In this chapter we seek to minimise the most familiar measure of dependence between two random variables, the Pearson product-moment coefficient, defined in Equation 4.14 (where the absolute value component was added to correctly represent the objective of our MNP model). The Pearson correlation coefficient between two variables is defined as the covariance of the two variables divided by the product of their standard deviations.

The Pearson coefficient is known to have some limitations. For example, being a measure of linear association between two variables, it may fail to capture a nonlinear relationship. Also, the coefficient can be drastically influenced by a few extreme outliers. In general, the Pearson coefficient is not considered suitable for non-normal distributed random variables. Egan (2007), for example, analysed the returns of the S&P500 index and concluded that the normal and lognormal distributions are a poor fit for its returns, the best fit being a t-distribution with location/scale parameters. The disadvantages of the Pearson coefficient are discussed in Joe (1997) and Hutchinson & Lai (1990).

There are alternative approaches that are considered more sensitive to nonlinear relationships, and which could be used as our objective function instead. The most popular alternative is the Spearman rank correlation coefficient ρ. The Spearman ρ is a "quasi-ordinal" correlation coefficient which is equivalent to the Pearson coefficient after the variables have been transformed into rank orders. The Spearman coefficient is often described as being "nonparametric" as a perfect Spearman correlation is obtained when two variables are related by any monotonic function, in contrast with the Pearson correlation, which only gives a perfect value when the variables are related by a linear function.

Another variable dependence coefficient that was introduced to address some of Pearson deficiencies is the Distance correlation, introduced by Székely et al. (2007) and Székely & Rizzo (2009). The Distance coefficient is analogous to the Pearson coefficient, but it is based on the pairwise Euclidean distances of each vector of random variables. This coefficient requires the calculation of the distance variance, distance standard deviation and distance covariance, defined similarly. The Distance correlation coefficient $\mathcal{R}$ satisfies $0 \leq \mathcal{R} \leq 1$, and an important property is that $\mathcal{R} = 0$ if and only if the random variables are statistically independent.

We leave the investigation of these alternative correlation measures for future work.

4.4.3 Alternative approaches to MNP

In this work we considered a MNP as a portfolio that is *uncorrelated* to the market benchmark, in which we considered the underlying index as the sole representation of the market. An alternative definition is to consider the market as represented by multiple factors, such as the Fama-French three-factor model (Fama & French (1993)) and the Carhart four-factor model (Carhart (1997)). If we redefine a MNP as a portfolio that is "independent" from multiple factors (instead of a single one), to retain a similar objective function we would need to transform Equation (4.14) into a function that minimises the multiple correlation coefficient (which takes values between zero and one), a much more complex and impractical task. An alternative is to redefine a MNP in the form of a "factor immune/independent portfolio" as defined below.

Under a factor assumption with M factors the standard (linear) equation for the return r_{it} on asset i at time t is:

$$r_{it} = \alpha_i + \sum_{j=1}^{M} \beta_{ij} F_{jt} + \text{ some random noise element} \tag{4.30}$$

In other words the return on asset i at time t is made up from some asset dependent term α_i plus factor terms made up from a linear sum of the factors, where β_{ij} is the coefficient for asset i in relation to factor j. Here the coefficients α_i and β_{ij} are time independent and are typically estimated from in-sample data using multiple least-squares regression.

The basic approach is therefore as follows. Using multiple least-squares regression estimate the coefficients α_i and β_{ij} in the in-sample period $[1, \ldots, T]$. Let $\hat{\alpha}_i$ and $\hat{\beta}_{ij}$ be the estimates. If we have a weight w_i associated with investment in asset i (where $\sum_{i=1}^{N} w_i = 1$) we approximate portfolio return at time t by the weighted sum of asset returns, namely $\sum_{i=1}^{N} w_i r_{it}$.

Turning to the factor equation and neglecting the noise term, it follows that the return on the portfolio at time t is given by:

$$\sum_{i}^{N} w_i r_{it} = \sum_{i=1}^{N} w_i \Big(\hat{\alpha}_i + \sum_{j=1}^{M} \hat{\beta}_{ij} F_{jt} \Big) = \sum_{i=1}^{N} w_i \hat{\alpha}_i + \sum_{j=1}^{M} \Big(\sum_{i=1}^{N} w_i \hat{\beta}_{ij} \Big) F_{jt} \qquad (4.31)$$

So under this equation the portfolio return (at time t) is composed of two terms:

$\sum_{i=1}^{N} w_i \hat{\alpha}_i$	a term dependent only on the assets in the portfolio
$\sum_{j=1}^{M} \Big(\sum_{i=1}^{N} w_i \hat{\beta}_{ij} \Big) F_{jt}$	a term involving time dependent factors

We want to minimise dependence on the market, where the market is driving the factors that we observe. Hence we want the influence of the factor term on the portfolio return to be as small as possible.

One way to achieve this is simply to minimise $|\sum_{t=1}^{T} \sum_{j=1}^{M} \Big(\sum_{i=1}^{N} w_i \hat{\beta}_{ij} \Big) F_{jt}|$, so the total factor contribution to portfolio return (summed over all time periods) is as small as possible. Other objectives are also possible, e.g. minimise $\sum_{t=1}^{T} |\sum_{j=1}^{M} \Big(\sum_{i=1}^{N} w_i \hat{\beta}_{ij} \Big) F_{jt}|$, which minimises the sum of the absolute values of factor terms at each time period t.

In fact, we did perform some preliminary investigations with this model and tested it with publicly available Fama-French factors for the US market, for portfolios composed of assets from the S&P500 index. The results we obtained were not encouraging and thus we did not do further research on this model.

4.5 Conclusions

In this chapter we considered the problem of constructing a market neutral portfolio where we can hold both long and short positions in assets. We formulated this problem as a mixed-integer nonlinear program, minimising the absolute value of the correlation between portfolio return and index return, and solved it using the Minotaur software package.

Computational results were presented for eleven different problem instances derived from universes defined by S&P international equity indices. These indicated that in-sample we could achieve very low correlations (in many cases zero correlation) in reasonable computation times. Out-of-sample correlations were higher, but for the majority of cases examined the market neutral portfolios constructed using the approach given in this chapter outperformed their benchmark indices.

Computational results, for the test problems considered, indicated that the model proposed out-performed an alternative approach based on minimising the absolute value of regression slope (the zero-beta approach).

We compared our approach with the performance of seven funds that adopt market neutral strategies with respect to the S&P 500. This comparison indicated that for three of these seven funds we had similar correlations, the other four funds had lower correlations than our market neutral portfolios. However in contrast to these seven funds (only one of which outperformed the index, and then only slightly) our market neutral portfolios outperformed the index by a significant amount in the vast majority of the cases examined.

Chapter 5

Exchange-traded funds: a survey and performance analysis

5.1 Introduction

ETFs (exchange-traded funds) have grown significantly in recent years, in terms of the number and size of funds and in trading volume. At their simplest ETFs offer replication of a market index such as the S&P500, and thereby offer the investor exposure to a market index in a much more flexible manner than a conventional mutual fund. In some countries ETFs also offer tax advantages over mutual or index funds.

Estimates of the total size of the ETF market vary, but as an indication BlackRock (2011) estimates it was approximately US$1.5 trillion (i.e. US$1.5 $\times 10^{12}$) at the end of 2011. The market size has doubled since late 2008. Due to the growth of the market for ETFs, regulators around the world have become concerned at their potential for inducing (or exacerbating) market risk and instability.

In the light of the growing importance of ETFs, we survey and classify existing ETFs and analyse their performance in replicating the behaviour of their underlying assets. We were able to identify 8192 ETFs (of which some are no longer active); we were able to find sufficient information to classify 6937 active ETFs. We selected a subset of 822 ETFs to conduct a detailed statistical performance analysis.

This chapter is structured as follows. We first discuss how plain vanilla ETFs and synthetic (leveraged/inverse) ETFs are constructed, mention regulatory concerns with regard to ETFs and review the academic literature relating to ETFs. We then describe our ETF survey database and generate insights into the ETF market by classifying this data (involving 6937 ETFs). The performance analysis of 822 ETFs follows.

5.2 ETF construction and literature review

Here we first discuss how ETFs are constructed and then briefly discuss concerns that have been expressed by regulatory authorities worldwide with respect to ETFs. We then go on to survey the literature with respect to ETFs.

5.2.1 ETF construction

Suppose that we wish to construct an ETF that replicates the performance of the S&P500 index and further suppose that we have \$1m in cash that we can use to create the ETF. The ETF creator (also known as a provider) first uses their cash to purchase a basket of all the stocks that are in the S&P500 in the same proportions as they are represented in the index. This basket will rise and fall in value exactly in line with the S&P500. Now the ETF creator can issue shares entitling the holder to a share of the underlying basket. If they issue 100000 (say) shares then as the underlying basket is worth \$1m each share is initially worth \$1m/100000 = \$10 (for simplicity, we ignore transaction costs associated with purchasing the basket of S&P500 stocks). As the value of the underlying basket changes (exactly in line with the S&P500) so too does the value of each share. In this manner the ETF share has an underlying value over time that exactly replicates the returns given by the S&P500.

Once the ETF creator has issued shares then they can be traded in the market. As they are traded their price may deviate, because of supply and demand considerations, from the underlying net asset value, NAV, to which the ETF shareholder is entitled. Here the NAV is defined as the total value of the underlying basket divided by the total number of ETF shares issued. However, any difference between an ETF share price and the underlying NAV gives rise to arbitrage possibilities. Hence ETF prices will (in practice) be arbitraged back to their underlying NAV (e.g. see Engle & Sarkar (2006), Kayali & Ozkan (2007) and Ackert & Yisong (2008) for a discussion of the evidence supporting the hypothesis that ETFs trade close to their underlying NAV).

One point to note here is that the ETF creator may not trade directly with all investors, instead the creator may trade directly only with a (small) selection of "authorised participants" who in turn trade with all investors. Authorised participants act as the interface between the ETF creator and the marketplace. One role these authorised participants play is to introduce liquidity into the market (i.e. they act as a market maker), by creating and redeeming ETF shares. They may create more ETF shares, by giving the ETF creator cash and/or appropriate stocks that they can use to enlarge the size (value) of the underlying basket. Alternatively, they may redeem ETF shares with the ETF creator for their underlying NAV, receiving cash and/or an appropriate proportion of the stocks in the basket. To exclude small trades, it is common for the ETF creator to constrain authorised participants to ensure that creations/redemption only take place in multiples of a given number of ETF shares (known as the “creation unit”, for example 50000 ETF shares).

This description captures the ETF process for a plain vanilla product that replicates the performance of an index. This index may, as above, be an equity index, but it could also be a bond or commodity index, or a single commodity price. Note here that the ETF creator makes three basic decisions: the **benchmark index** to use, the **target return** (tied to the chosen index) and the **basket** to hold to achieve that return. Once these decisions have been made the success (or failure) of the ETF depends upon its ability to attract marketplace investors. Readers interested in greater detail as to the process of creating ETFs and the role played by authorised participants are referred to Gastineau (2004), Deville (2008), Gastineau (2010), IndexUniverse (2011) and Investment Company Institute (2012).

Historically ETFs started out as index trackers, aiming to give the same return as a benchmark index. Here, because index composition is known, the ETF basket can fully replicate the index. Alternatively, approaches based on replicating the index by holding a subset of the assets in the index could be used to decide the composition of the basket (e.g. see Beasley et al. (2003) and Canakgoz & Beasley (2009) for approaches to index tracking).

As ETFs evolved, their scope widened beyond index tracking. Leveraged ETFs, aiming to give a multiple of index return (e.g. $1.5\times$ or $2\times$), appeared. Inverse (or short) ETFs aiming to give the negative of index return (so $-1\times$) also appeared. Here we refer to an ETF as a $L\times$ ETF if the ETF aims to return a multiple L of the benchmark index return. So an index tracking ETF would be referred to as a $1\times$ ETF. ETFs aiming to give a leveraged inverse (say $-1.5\times$ or $-2\times$) have also now appeared. Note here that $L\times$ ETFs with $L > 1$ are sometimes called “bull ETF” $L\times$ ETFs with $L \leq -1$ are sometimes called “bear ETFs”.

For leveraged and inverse ETFs deciding the composition of the basket that the ETF creator should hold so as to achieve the target return is a genuinely difficult task. Consequently “synthetic” ETFs have been developed. For ETFs of this type the ETF creator enters into a contract with a third-party, typically a bank. At its simplest this third-party invests on behalf of the ETF creator and promises to deliver the target return of $L\times$ with respect to the benchmark index specified. The advantage of this arrangement from the viewpoint of the ETF creator is that it relegates all issues relating to basket composition to a third-party. Indeed the ETF creator does not even have to know the details of how the third-party delivers $L\times$, merely needs to trust that the third-party will deliver their promised return of $L\times$. In practice third-parties typically employ swaps/futures/derivative contracts to deliver the promised return. Leveraged/inverse ETFs require daily rebalancing in order to achieve promised returns. Cheng & Madhavan (2009), Rollenhagen (2009) and Little (2010) discuss leveraged and inverse ETFs in greater detail.

A more recent evolution in the ETF market has been the introduction of actively managed ETFs where professional managers actively trade the underlying basket in an attempt to generate return. In ETFs of this type there is (potentially) no need for the ETF to be tied to an index.

5.2.2 Regulatory concerns

As mentioned above due to the growth of the market for ETFs, regulators around the world have become concerned at their potential for inducing (or exacerbating) market risk and instability. In the USA a US Senate (2010) report into the "flash crash" of May 6th 2010 discussed the role that ETFs played in that market event and in October 2011 the US Senate Committee on Banking, Housing and Urban Affairs conducted a hearing into ETFs. One point arising from this US Senate investigation is that when ETF market makers detect unusual movements that raise questions about the price of one or more of the stocks in the basket underlying an equity based ETF then they may cease to provide liquidity in that ETF. Madhavan (2012) provides an example of a particular ETF whose price, although usually closely tracking the intraday NAV based on the underlying basket, deviated significantly from that NAV for 25 minutes during the flash crash.

In the UK, in June 2011, the Bank of England Financial Stability Report (2011) noted that the growth and evolution of the ETF market has sparked regulatory concern. That report identifies three issues: complexity, interconnectedness and opacity associated with financial instruments and notes that synthetic (e.g. leveraged/inverse) ETFs which lack an underlying basket of physical/equity assets, are a particular cause for concern in relation to systemic risk. In June 2012 the Bank of England Financial Stability Report (2012) rated 18 different financial instruments and synthetic ETFs were one of only four judged to be in the worst category of "highly opaque".

Internationally in April 2011, the Financial Stability Board (2011) (national regulators from 24 countries, including the G-20) highlighted the potential stability risks arising from ETFs. Also in April 2011, Ramaswamy (2011), writing for the Bank of International Settlements, identified a number of channels by which ETFs could contribute to financial stability risk. Complexity, opacity, synthetic ETFs and liquidity were some of the issues highlighted there, reflecting the fact that broadly speaking regulators worldwide have similar concerns about ETFs.

5.2.3 Literature review

To summarise the relevant literature relating to ETFs, we have focused primarily on published academic literature. ETFs were introduced in the 1990s, some early issues around their introduction are discussed in Kupiec (1990) and Gastineau (2001). Poterba & Shoven (2002) provide some statistics on the growth of ETFs since their introduction in the 1990s. The total ETF market was approximately US$80bn in 2001 with just two ETFs (the SPDR Trust SPY and the Nasdaq-100 QQQ) making up some 60% of the market. Boehmer & Boehmer (2003) considered the introduction by the New York Stock Exchange (NYSE) of trading in three large ETFs (SPY, QQQ and a Dow Jones ETF, DIA), plus a number of smaller ETFs, that had previously been traded just on other exchanges. They concluded that this introduction had lowered ETF trading costs. Kostovetsky (2003) examined the conditions under which it is preferable for an investor to invest in an (index tracking) ETF as compared with a conventional index tracking mutual fund. Alexander & Barbosa (2008) examined the hedging problem which arises in ETF creation/redemption when the basket underlying the ETF shares involves illiquid stocks with relatively high transaction costs.

Mariani et al. (2009) examined the return distributions of three ETFs and their corresponding benchmark indices using a Levy model, concluding that the ETFs exhibit the same behaviour as their respective indices. Avellaneda & Zhang (2010), Giese (2010) and Jarrow (2010) presented models for a leveraged (and inverse) ETF by assuming that the ETF follows a diffusion process. Lin & Mackintosh (2010) discussed issues related to tracking error calculations for ETFs. With respect to papers that focus on ETF performance we present Table 5.1, where we summarise the scope of each study (ETFs covered) and its conclusions.

Table 5.1: Papers dealing with the performance of ETFs

Paper	Scope	Conclusions
Elton et al. (2002)	One index tracking ETF over the period 1993 to 1998	ETF underperformed the index by 28.4 basis points
Poterba & Shoven (2002)	One index tracking ETF over 7 years	ETF returned 19.17% per year, a mutual fund 19.33%, the index 19.39%.
Agrrawal & Clark (2007)	38 ETFs over the period 2002 to 2007	Regression of ETF return against market return indicates that regression slopes (beta's) are not affected by return frequency, but are affected by estimation period
De Jong & Rhee (2008)	Up to 217 ETFs with weekly data over the period 1996 to 2005	Momentum and contrarian strategies will yield significant excess returns
Hlawitschka & Tucker (2008)	9 ETFs over the period 2002 to 2005	Performance of a mean/variance portfolio drawn from the major stock constituents of the ETFs superior to other portfolios examined

Continued on next page

Table 5.1 – continued from previous page

Paper	Scope	Conclusions
Rompotis (2008)	62 German ETFs with weekly data over the period 2000 to 2006	ETFs slightly underperform their benchmark indices; they have greater risk (standard deviation in return) than their indices
Elston & Choi (2009)	One $-1\times$ ETF and five $-2\times$ ETFs in 2008	5 of the 6 ETFs underperformed with respect to their target return
Johnson (2009)	Daily and monthly data for 20 index tracking ETFs over the period 1997 to 2006	Explanatory factors for the correlations found included return relative to USA index and overlapping exchange opening hours
Maister et al. (2009)	505 US-listed ETFs in 2008	Average difference between NAV return and index return was 52 basis points
Rompotis (2009)	20 index tracking ETFs within the period 2004 to 2006	ETFs slightly underperform their benchmark indices
Aber et al. (2009)	4 index tracking ETFs	Mixed picture as to tracking ability with respect to returns achieved
Maister et al. (2010)	563 US-listed ETFs in 2009	Average difference between NAV return and index return was 125 basis points
Shin & Soydemir (2010)	26 ETFs with daily data over the period 2004 to 2007	ETFs underperform their benchmark indices
Wong & Shum (2010)	Daily performance of 15 ETFs over the period 1999 to 2007	ETF returns are higher in bullish, than bearish, markets; some ETFs with the same benchmark index perform differently
Agapova (2011)	Monthly performance of 11 ETFs with comparable index tracking funds over the period 2000-2004	Very few significant differences between ETFs and index tracking funds
Charupat & Miu (2011)	8 Canadian leveraged ($2\times$, $-2\times$) ETFs, compared with four non-leveraged ($1\times$, $-1\times$) ETFs	Leveraged ETFs more actively traded than non-leveraged ETFs; daily returns regression indicated that the ETFs were giving returns close to the $\pm2\times$ promised
Rompotis (2011a)	37 inverse leveraged ($-2\times$, $-3\times$) ETFs within the period 2006 to 2011	ETFs underperform their daily target return
Rompotis (2011b)	50 index tracking ETFs over the period 2002 to 2007	ETFs outperformed the S&P500; tracking error with respect to the benchmark index is strongly persistent in the short term
Rompotis (2011c)	14 actively managed ETFs within the period 2008 to 2010	No significant difference with regard to average daily return and risk when comparing the ETFs to the S&P500
Sabbaghi (2011)	15 green ETFs within the period from 2005 to 2009	Positive cumulative returns from inception through to end of 2008, negative thereafter
Schmidhammer et al. (2011)	Five ETFs and three index certificates replicating the DAX using minute prices over two months in 2008	ETFs based on complete replication perform better than index certificates or ETFs based on swaps
Blitz & Huij (2012)	7 global emerging markets equity ETFs from inception until December 2010	High levels of tracking error, higher than developed market ETFs
Blitz et al. (2012)	3 European ETFs over the period 2003 to 2008	Dividend taxes and expense ratios contribute to underperformance
Buetow & Henderson (2012)	845 US-listed ETFs over the period 1994-2010	On average ETFs closely track their benchmark index
Haga & Lindset (2012)	Four Norwegian leveraged ($2\times$, $-2\times$) ETFs over the period from January 2008 to May 2010	Regression indicated that the ETFs are not achieving the $\pm2\times$ returns promised

Continued on next page

Table 5.1 – continued from previous page

Paper	Scope	Conclusions
Rompotis (2012)	68 leveraged and inverse (2×, −1×, −2×, −3×) ETFs	ETFs are not achieving the returns promised, on average the majority of daily returns deviate from the target multiple by at least 10 basis points
Sharifzadeh & Hojat (2012)	34 ETFs, matched with passive index mutual funds, over the period 2002 to 2010	No statistical support for the hypothesis that ETFs outperform index funds; no overall difference between ETFs and index funds in terms of Sharpe ratio

The 27 papers seen in Table 5.1 cover a range of time periods and a range of ETFs. Most studies involve only a small number of ETFs and concentrate solely on a limited range of ETFs (so only consider $L\times$ ETFs for one or two different values of L). Of the three studies with a large number of ETFs, two studies (Maister et al. (2009, 2010)), each cover only a single year. The study by Buetow & Henderson (2012) does cover multiple years, but does not separately identify ETFs of varying types $L\times$. In view of the limitations of these analyses, we consider that a much fuller analysis is necessary. Our analysis covers a large number of ETFs over many years and considers different values of L (i.e. looks at leveraged and inverse ETFs). Our analysis provides a comprehensive snapshot of the ETF market with 6937 ETFs being surveyed and classified; further, it presents a detailed statistical performance analysis relating to 822 ETFs, of varying types $L\times$ ($L = +1, -1, +2, -2, +3, -3$), over all years from 1993 onwards.

5.3 ETF survey

In our survey of the ETF market the information given here was collected from Thomson Reuters Datastream (2013) and is a market snapshot taken in September 2011. According to Datastream, in September 2011, there were a total of 7198 active and 994 dead or suspended ETFs, giving a total of 8192 ETFs. Datastream does not include information on which benchmark indices (if any) are tracked by each ETF. For this reason we had to manually carry out extensive work to find this information for as many ETFs as possible, as well as to classify ETFs by their target return (i.e. find the value of L for each $L\times$ ETF). Much of this manual work required examination of individual ETF websites.

Although the ETF market started in 1993, it has experienced a sharp increase in the past 5-6 years. Figure 5.1 shows the cumulative number of ETFs created over the years, (including those that are currently dead or suspended). A sharp rise can be seen from 2005 onward. Since then the number of ETFs created has soared, from less than 1000, to 8192 as of September 2011. From the end of 2005 to the end of 2010 (the last full year for which we have data) the number of ETFs increased at a compound rate of 55% per year. In the 12 months to September 2011, the date of our snapshot, 1578 new ETFs were launched, a creation rate of over 6 ETFs per trading day. Of these 8192 ETFs, 902 are leveraged/inverse ETFs. There are 2018 different underlying benchmarks associated with these 8192 ETFs.

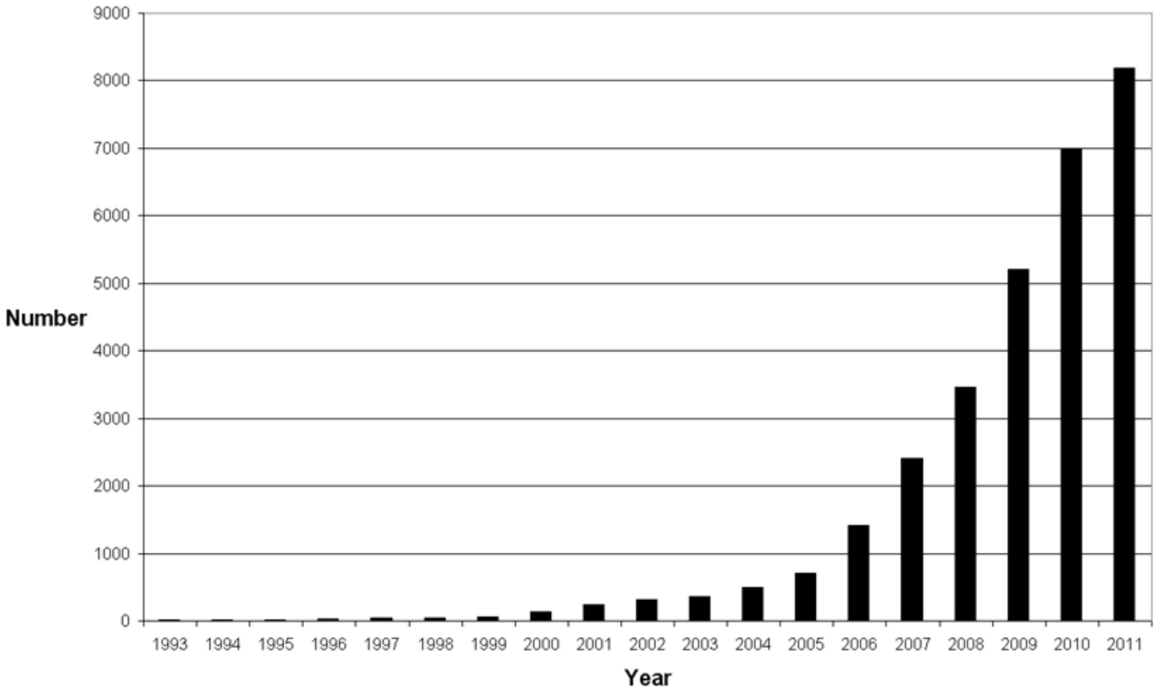

Figure 5.1: Cumulative number of ETFs over time

All ETFs were classified into one of several major categories (e.g. Single Market Equity Tracker), and then further subdivided within those major categories (e.g. Real Estate Sector). The results of this classification are shown in Table 5.2; the first column shows the major classification, the second column the sub-classification; within each sub-classification the number of active ETFs for each desired return (e.g. $L\times$) is given; the final two columns give the total number of active and dead/suspended ETFs within that sub-classification.

As can be seen from Table 5.2 the vast majority of active ETFs track equity indices, 2607 ETFs (37.6% of active ETFs) track single market equity indices, 2272 (32.8%) multi-market equity indices. The next most common categories are commodity (13.5%) and bond (12.4%) trackers. In terms of the type of performance expected, 87.7% of active ETFs are simple trackers ($1\times$); 4.4% are inverse ETFs ($-1\times$), so offering the equivalent of shorting the underlying benchmark; 4.3% are leveraged ($2\times, 3\times$); 2.5% are inverse leveraged ($-2\times, -3\times$); 0.8% offer excess return.

In Table 5.3 we show the creation date of the (currently) active ETFs as seen in Table 5.2, subdivided by category. 514 of these ETFs were created before the end of 2005. The

Table 5.2: ETF summary by number within each major and sub-classification and performance type

Major classification (no. active)	**Sub-Classification**	**1×**	**−1×**	**2×**	**−2×**	**3×**	**−3×**	**Exc. Return**	**Other**	**Active**	**Dead**
Bond Tracker (860)	Bond Index	785	41	5	6	2	2	6		847	102
	Bonds	1	2	5	5					13	1
Commodity Tracker (935)	Commodity	244	15	18	7			4	4	292	7
	Commodity Futures	7	1	2	3			13		26	8
	Commodity Index	428	40	74	2			31		575	0
	Commodity Index 3 month forward	42								42	0
Currency Tracker (202)	Currency	66	42	3	3	14	14			142	4
	Currency Index	42	1							43	0
	Interest Rate Index	17								17	4
Derivative Tracker (2)	General	2								2	1
Hedge Fund Tracker (26)	Hedge Fund Index	26								26	11
Inflation Tracker (2)	Inflation Index							1		1	0
	Interest Rate Index	1								1	0
Interest Rate Tracker (1)	Interest Rate Index		1							1	0
Loan Market Tracker (1)	Loans	1								1	0
Multi-Market Equity Tracker (2272)	Futures	5								5	0
	General	1215	26	40	33	3	3	1		1321	159
	Real Estate Sector	66								66	8
	Sector	760	44	4	4	1	1			814	118
	Specialised	66								66	17
Multi-Asset Index Tracker (16)	Comm. or Bond or Equity / Currency Index	1							5	6	0
	General	10								10	0
Real Estate Tracker (9)	Mortgage	4								4	2
	Real Estate Index	5								5	0
Single Market Equity Tracker (2607)	Futures	10		1						11	1
	General	1763	86	92	59	10	8		2	2020	210
	Real Estate Sector	49	1	1	2	1	2			56	10
	Sector	416	8	20	17	7	6			474	106
	Specialised	46								46	22
Volatility Tracker (4)	Implied Volatility Index	2								2	0
	Volatility Index	2								2	0
Total		**6082**	**308**	**265**	**141**	**38**	**36**	**56**	**11**	**6937**	**791**
Percentage of active ETFs		**87.7**	**4.4**	**3.8**	**2**	**0.5**	**0.5**	**0.8**	**0.2**		

Notes:

(a) A multi-market equity tracker means that the ETF benchmark index contains stocks from two or more national/country markets (for example the S&P1200)

(b) A single market equity tracker means that the ETF benchmark index contains stocks from only one national/country market (for example the FTSE100)

(c) Real estate trackers track property price indices, in equity indices the real estate sector trackers follow indices related to stocks in real estate companies

(d) Specialised equity trackers track indices that are subject to specific policies, such as Islamic indices

(e) For bond, commodity and currency trackers some track indices and others track prices (or future prices)

(f) There are 7728 ETFs in the above table, we were unable to find sufficient information to accurately classify 464 (approx. 5.7%) of the 8192 ETFs in our survey (recall that this classification is not automatic in Datastream but must be manually generated, e.g. by individually examining each ETF website) vast majority of non-equity ETFs was formed after 2005. The first of these ETFs to offer excess/leveraged return was launched in 2005. Inverse and inverse leveraged ETFs began appearing in 2006.

Table 5.3: Number of active ETFs by year of introduction and category

Year	**Single Market Equity Tracker**	**Multi-Market Equity Tracker**	**Commodity Tracker**	**Bond Tracker**	**Currency Tracker**	**Hedge Fund Tracker**	**Multi-Asset Index Tracker**	**Real Estate Tracker**	**Others**	**Total**
1993	1									**1**
1994										
1995	2									**2**
1996	18									**18**
1997	2									**2**
1998	9	2								**11**
1999	5	1								**6**
2000	47	7		2						**56**
2001	46	15						1		**62**
2002	25	11		4						**40**
2003	19	10	1	9						**39**
2004	62	48	4	12						**126**
2005	79	53	7	11	1					**151**
2006	234	174	92	55	7		2	1	1	**566**
2007	287	272	151	98	7		2	1	1	**819**
2008	293	340	135	64	34		5	1		**872**
2009	401	462	182	249	53	7		3	3	**1360**
2010	663	509	260	194	62	7			1	**1696**
2011	414	368	103	162	38	12	7	2	4	**1110**
Total	**2607**	**2272**	**935**	**860**	**202**	**26**	**16**	**9**	**10**	**6937**

Note: 2011 data ranges from January to September

Considering Table 5.3 we can see that, even though equity trackers still dominate the ETF market, there has been clear diversification in recent years. If we take the ETF market as it was at the end of 2005, equity trackers constituted (by number) 90% of the entire market. By the end of 2007, this had fallen to 75%, falling further to 70% now. This fall in equity ETFs has been compensated for by a rise in bond and commodity ETFs which were jointly responsible for just 10% of the entire ETF market at the end of 2005, 23% by the end of 2007 and, finally, 26% at the time of our market snapshot.

Considering both Figure 5.1 and Table 5.3 it would be hard to discern that in 2007-8 there was a global financial crisis. Even looking within the data for any sign of a shift in ETF emphasis it is hard to identify any significant effect. For example one might hypothesise that the financial crisis would divert investment attention from older economies to newer, and more emerging, economies. At the end of 2007, before the financial crisis had taken hold, 79% of all single market ETFs tracked markets in Europe, North America (USA/Canada) or Japan. By the end of 2009, two years later and after the financial crisis had hit, 77% tracked these markets. This 2009 percentage is effectively the same as the 2007 percentage, despite the growth in ETFs that occurred in the meantime, as evidenced in Table 5.3. Of the 545 single market ETFs started in the 12 months to September 2011 64% tracked markets in Europe, North America or Japan. Overall these data do not appear to indicate that the global financial crisis has diverted the ETF market away from older economies.

In terms of the size of each ETF we were able to get the market value (equivalent to the number of ETF shares times the NAV) for approximately 30% of active ETFs. Information from Thomson Reuters Datastream (2011) indicated that they rely on ETF providers to supply this information and some do not. It was clear from our data that lack of market value information was more of an issue with newer ETFs than older ETFs. For example we had market values for 65% of the ETFs created before 2006; for ETFs created in 2010 we had market values for just 21%.

The available market values (MVs) are summarised in Table 5.4, all converted into US$ for ease of comparison. Over half the market value is in single market equity trackers; nearly 70% of market value is in equities in some form. We can see 20.9% of market value is in commodity trackers despite being only 7.4% of active ETFs by number. Commodity trackers have the highest (arithmetic) mean market value; this is contrast with multi-market equity trackers which have a comparatively low mean market value. Comparing the (arithmetic) mean and median ETF size (as in the ratio column in Table 5.4) reveals that the distribution of ETF MVs within each sector is highly skewed. This is most clearly apparent for commodity trackers, where the ratio is 59.7 (so the mean ETF MV is nearly 60 times larger than the median ETF MV). In fact in this category 95% of the total MV is concentrated in 10% of the ETFs by number.

Table 5.4: ETF market value (MV) summary

Classification	**Number active**	**Number with available MV**	**% with available MV**	**Total MV (US$m)**	**MV % total**	**Mean ETF MV (US$m)**	**Median ETF MV (US$m)**	**Ratio (mean/ median)**
Bond Tracker	860	252	11.9	238167.7	8.1	945.1	80.7	11.7
Commodity Tracker	935	157	7.4	618330.9	20.9	3938.4	66	59.7
Currency Tracker	202	42	2	55296.2	1.9	1316.6	112.2	11.7
Derivative Tracker	2	2	0.1	148.8	0	74.4	74.4	1
Hedge Fund Tracker	26	5	0.2	444	0	88.8	29.7	3
Inflation Tracker	2	1	0	26	0	26	26	1
Loan Market Tracker	1	1	0	166.4	0	166.4	166.4	1
Multi-Market Equity Tracker	2272	473	22.3	484177.2	16.4	1023.6	58.7	17.4
Multi-Asset Index Tracker	16	14	0.7	1492.7	0.1	106.6	77.7	1.4
Real Estate Tracker	9	7	0.3	6044.9	0.2	863.6	50.9	17
Single Market Equity Tracker	2607	1171	55.1	1551126.2	52.5	1324.6	63.7	20.8
Total	**6932**	**2125**		**2955421**				

One point of interest from Table 5.4 relates to the total size of the ETF market.

BlackRock (2011) estimated the size of the ETF market as approximately US$1.5 trillion (i.e. US$$1.5 \times 10^{12}$) at the end of 2011, and involving (at most) 4200 ETFs. Although the precise classification of a particular fund as an ETF can vary, as do daily market values, our survey indicates that these figures potentially underestimate the total size of the market. In Table 5.4, which uses data for 2125 ETFs (31% of 6932 active ETFs) we find a total market value of US$2.96 trillion (so approximately US$3 trillion). As with commodity trackers, there is a distinct Pareto effect in ETF market values: 13% of ETFs represent 90% of market value; 7% of ETFs make up 80% of total market value. In fact the ETF market is so highly skewed that just 28 ETFs make up 50% of total market value.

We now look at the larger categories of ETFs in more detail. Firstly, in Figure 5.2, we consider single market (country) equity ETFs, the performance of these ETFs is linked to an index in a particular country, either a market index or a more specialised sector index. The figure shows the top twenty countries ranked by ETF MVs, the number of ETFs is also shown. We can see that the United States dominates with 25% of total MV; ETFs following indices in China represent 17%, Japan 15%. The remaining BRIC countries (Brazil, Russia, India, excluding China) represent 7% of total MV. Considering ETF numbers, 40% follow the United States; roughly five times more than follow the next most popular single country, China.

Secondly, we consider equity ETFs following multi-market (country) indices. Figure 5.3 shows the top twenty (by MV) indices tracked. It can be seen that MV is highly concentrated, with 29% of MV associated with ETFs following emerging markets. EAFE (Europe, Australasia and Far East) countries account for 24% of MV; Europe accounts for 20% and ETFs following global indices account for 16%. With respect to the number of ETFs, 40% track European, and 25% global, indices.

Thirdly, we consider commodity based ETFs and in Figure 5.4, we show the top ten commodities or commodity indices tracked. Again MV is highly concentrated; with ETFs tracking gold accounting for 53% of MV and 38% of MV is in ETFs tracking WTI (West Texas Intermediate) oil futures contracts. Crude oil is tracked by several other indices, but WTI dominates. The third largest commodity tracked by MV is silver with 5% of market value. Taken together these three commodities account for approximately 95% of total MV associated with commodity ETFs. In terms of the number of ETFs, 11% follow gold, 8% follow general commodity indices (Dow Jones - UBS Commodity index or the S&P GSCI), 6% follow platinum or palladium and 5% follow silver.

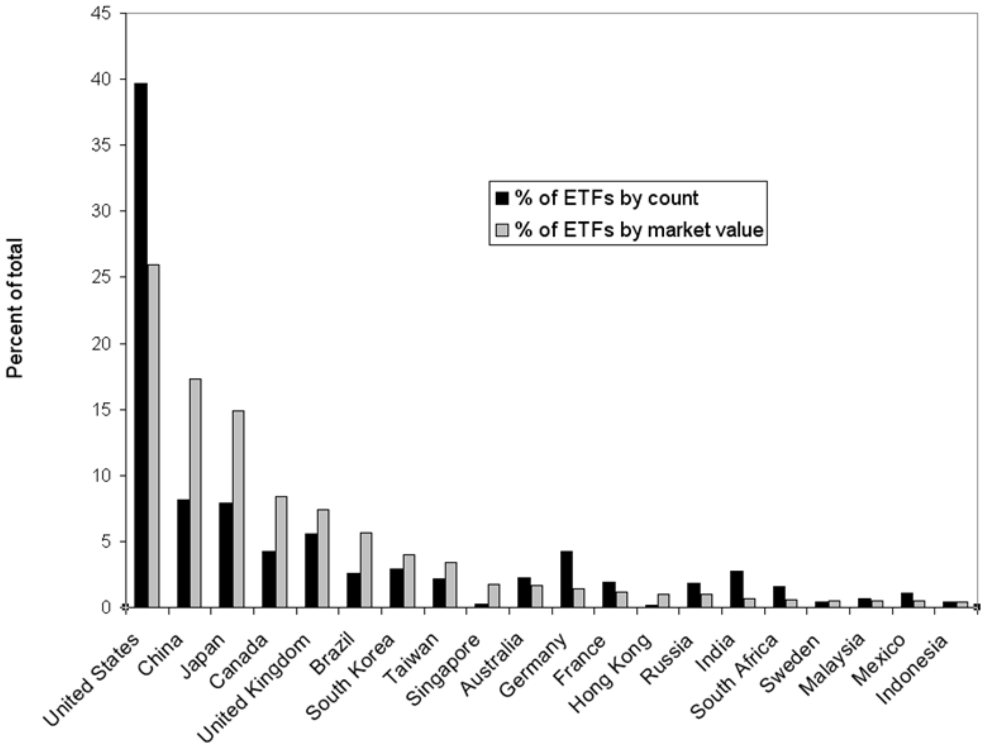

Figure 5.2: Single market equity ETFs: the top 20 countries by market value, showing percentage of total by market value and number of ETFs

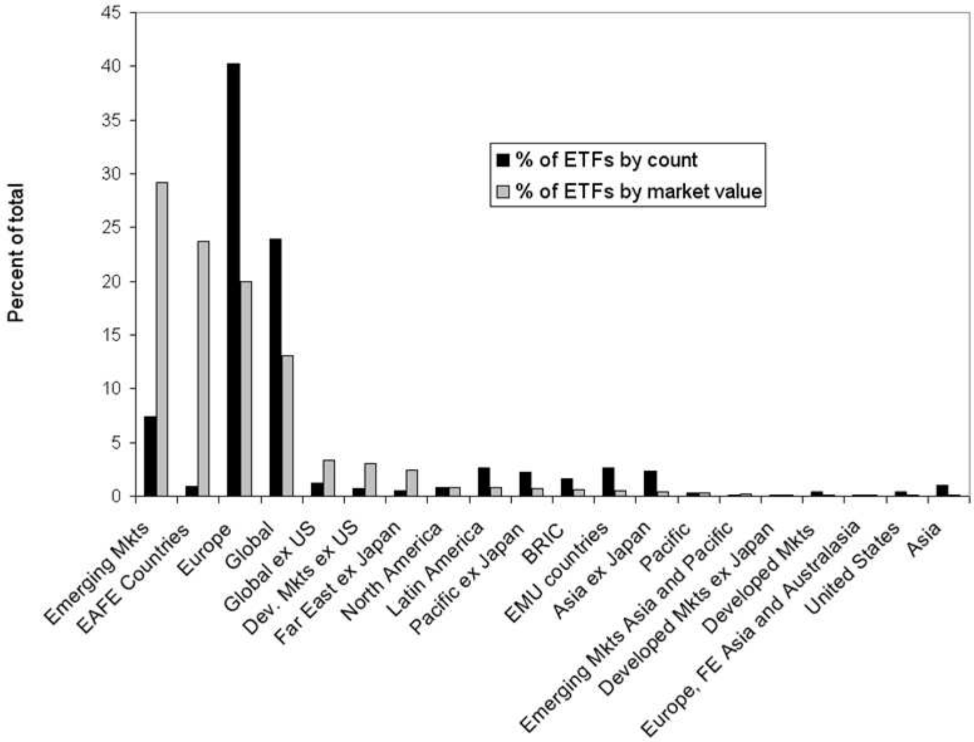

Figure 5.3: Multi-market equity ETFs: the top 20 indices by market value, showing percentage of total by market value and number of ETFs

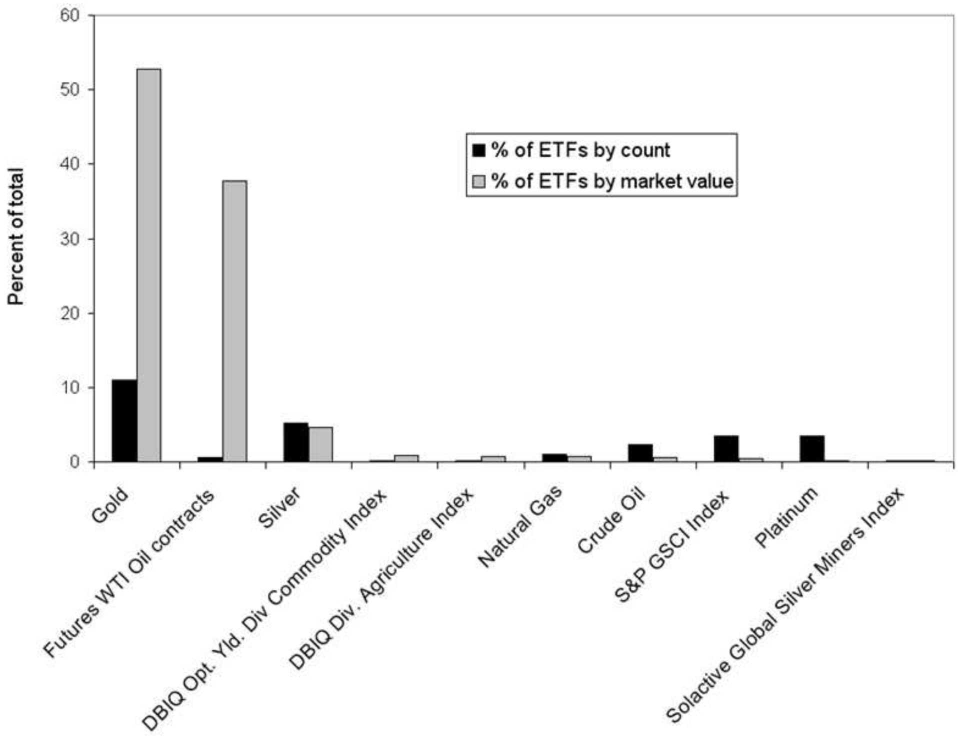

Figure 5.4: Commodity ETFs: the top 10 commodity or commodity indices by market value, showing percentage of total by market value and number of ETFs

5.4 ETF performance

The term performance is used here to denote the accuracy with which an ETF replicates the return behaviour of its benchmark. We first discuss how our database of ETFs for performance analysis was selected and then present the statistics we calculated.

We draw our performance database from the 2125 active ETFs for which we had a market value (recall from above that total MV was US$2.96 trillion). Since our snapshot is as at the end of September 2011 we excluded any ETFs that were created after September 2009 (so we had at least two years of data), and this left 1413 ETFs as potential candidates for analysis. For these candidates we collected daily price and benchmark index values from Datastream with which to calculate daily returns. The price series for some ETFs were intermittently reported so unless an ETF had at least 70% of possible return observations available it was not included in our performance database. A small number of ETFs were also excluded as a result of a preliminary analysis which indicated that they appeared to be outliers, probably due to a misinterpretation on our part as to the underlying benchmark index. Our performance database, after the process described above, contained 822 ETFs. This represents 39% of the initial 2125 ETFs, but these 822 ETFs had a total MV of US$1.81 trillion, so we captured in our database 61%

$(= 100(1.81/2.96)\%)$ of ETFs by MV. The ETF with the most return observations in this database, 4755 dating from 1993, was the very first ETF, the SPDR Trust SPY, which tracks the S&P500.

These ETFs were associated with 444 different benchmark indices. Note here that we explicitly identified the benchmark index (and the associated value of L) for each ETF in our performance database. Our performance database contained 686 tracking ($1\times$) ETFs, 16 inverse ($-1\times$) ETFs, 58 positive leveraged ETFs (51 $2\times$, 7 $3\times$) and 62 inverse leveraged ETFs (one $-1.5\times$, 55 $-2\times$, 6 $-3\times$) ETFs. In terms of our classification (as in Table 5.2) there were 59 bond trackers (all tracking an index); 57 commodity trackers (37 commodity, one futures and 19 index); 21 currency trackers (all tracking a currency); 2 derivative trackers; 204 multi-market equity trackers (139 general, 8 real estate sector, 55 sector, 2 specialised); 4 multi-asset index trackers (all general); 4 real estate trackers (2 mortgage, 2 real estate index); 471 single market equity trackers (322 general, 16 real estate sector, 130 sector, 3 specialised).

To perform our analysis, we used returns based on the daily changes of log prices (e.g. the return on an ETF at (trading) day t is $r_t = \ln(\text{ETF price at day } t/\text{ETF price at day } t-1))$. We define the benchmark return B_t on day t for a $L\times$ ETF using $B_t = L\times$ (ETF benchmark index return on day t), this allows us to compare $L\times$ ETFs with varying values for L (either positive or negative) in a consistent manner.

The investor holds an ETF with the expectation that its return behaviour will mimic that of the underlying benchmark index. This behaviour can be summarised by average return and volatility. Let us define:

$\mu_r = E(r_t)$ and $\sigma_r^2 = V(r_t)$ as the mean and variance of the ETF's return;
$\mu_B = E(B_t)$ and $\sigma_B^2 = V(B_t)$ as the mean and variance of the benchmark's return.

The difference in mean return is $(\mu_B - \mu_r)$ and the difference in volatility is measured by the difference in variances $(\sigma_B^2 - \sigma_r^2)$.

Let us model the return of the ETF as a linear model of the benchmark

$$r_t = \alpha + \beta B_t + \varepsilon_t \qquad \text{where } V(\varepsilon_t) = \sigma_\varepsilon^2. \tag{5.1}$$

Ideally the ETF perfectly reproduces the behaviour of the index and α is zero, β is unity and σ_ε^2 is zero. Underperformance occurs if $(\mu_B - \mu_r)$ is positive or (since $\mu_r = \alpha + \beta\mu_B$) if

$$(1-\beta)\mu_B - \alpha > 0. \tag{5.2}$$

Similarly $\sigma_r^2 = \beta^2\sigma_B^2 + \sigma_\varepsilon^2$, thus the ETF is more volatile than the benchmark $(\sigma_r^2 > \sigma_B^2)$

if

$$\sigma_\varepsilon^2 - (1-\beta^2)\sigma_B^2 > 0. \tag{5.3}$$

5.4.1 Underperformance in mean return

In Figure 5.5, we compare the mean return of each ETF with that of its benchmark using all the data available for each benchmark (shown as % p.a.). The diagonal line in Figure 5.5 divides the plot into two triangles. Points in the top left (upper) triangle of the plot show ETFs which produce a greater mean return than their benchmarks ($\mu_r > \mu_B$); in general this outperformance is small. Points in the lower right triangle show ETFs whose mean return is less than that of their benchmarks; there are many examples of severe underperformance by ETFs. The distribution of $(\mu_B - \mu_r)$ is positively skewed: the lower quartile is -0.51% p.a.; the median is 0.59% p.a.; the upper quartile is 3.9% p.a.; the 95 percentile is 33% p.a. Describing the variation in accuracy another way, 37% (53%) of ETFs yielded a mean return within 1% p.a. (2% p.a.) of the benchmark return.

In order to investigate why some ETFs fail to reproduce the benchmark return the possible causes of the failure we consider are the degree of (inverse) leverage required; the category of the benchmark; the periods over which the returns were observed. The following regression was estimated.

$$\begin{aligned}\mu_B - \mu_r = \quad & \theta_0 + \theta_1 L_{-2} + \theta_2 L_{-1} + \theta_3 L_2 + \theta_4 \ln(\text{Market Value}) \\ & + \theta_5 \text{Cat(Bond)} + \theta_6 \text{Cat(Commodity)} + \theta_7 \text{Cat(Currency)} \\ & + \theta_8 \text{Cat(Multi-market)} + \theta_9 (\text{pre } 2000) + \theta_{10} (\text{pre } 2005) \\ & + \theta_{11} (\text{pre } 2007) + \theta_{12} (\text{pre } 2009) + \text{error}\end{aligned} \tag{5.4}$$

where L_{-2} is a zero/one binary indicator that is unity if the ETF is inverse and leveraged by a factor of 2 or more (zero if not); L_{-1} indicates whether an ETF is inverse (non-leveraged), L_2 indicates whether an ETF is leveraged by a factor of 2 or more. Cat(Bond) is a binary indicator that is unity if the ETF is a Bond Tracker (zero if not); the other Cat() indicators are defined similarly. The binary indicators describing the start date of the ETF are set to unity if the start date is before the beginning of the year mentioned. If the ETF is a simple single equity market tracker with start date in February 2009, for example, then all the binary indicators are zero.

Note here that given regression equation (5.4) and the fact that underperformance occurs if $(\mu_B - \mu_r)$ is positive a negative coefficient (θ_1 to θ_{12}) indicates a factor that contributes to reducing underperformance, a positive coefficient a factor that increases underperformance. For all of the regression results given in this chapter we only regard a

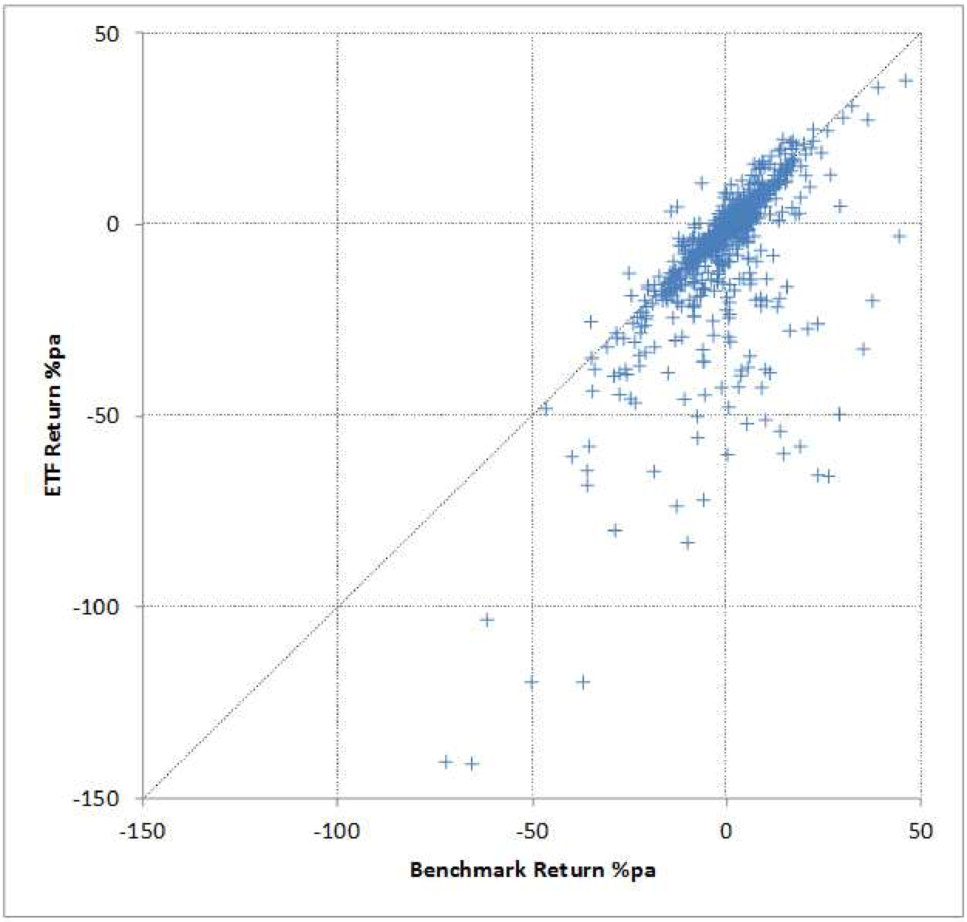

Figure 5.5: A comparison of ETF mean return with that of its benchmark. The mean is calculated over all available data

regression coefficient as significant if it has a p-value of 0.01 or less (so a 1% significance level).

The estimation results are summarised in Table 5.5, where the significant regression coefficients have been highlighted. The R^2 value is 66%, which means that over half the variation in $(\mu_B - \mu_r)$ can be explained by the properties of the ETFs. On average, both inverse trackers and leveraged trackers have significantly greater underperformance than a single equity tracker (since they have significant positive coefficients in Table 5.5). Currency trackers, with a significant negative coefficient in Table 5.5, tend to track their benchmark slightly better than a single equity tracker.

Looking at equation (5.2) underperformance can be decomposed into failing to fully capture the direction of changes in the benchmark, measured by $(1-\beta)\mu_B$, or a systematic failure to capture the level of returns, measured by $-\alpha$. The median (upper quartile) value of $(1-\beta)\mu_B$ is 0.00012% per day (0.0020% per day) whereas the median (upper quartile) value of $-\alpha$ is 0.0021% per day (0.0190% per day). It is clear that failure to capture the

Table 5.5: Regression to explain $(\mu_B - \mu_r)$ in terms of the properties of the ETF

Analysis of $(\mu_B - \mu_r)$	$R^2 = 0.66$		
	Coefficient	*Standard Error*	*p-value*
Intercept	-0.003	0.004	0.42
L_{-2}	***0.157***	***0.004***	***0.00***
L_{-1}	***0.052***	***0.008***	***0.00***
L_2	***0.054***	***0.005***	***0.00***
ln(MV)	0.000	0.001	0.97
Bond	0.007	0.004	0.10
Commodity	0.001	0.004	0.78
Currency	***-0.030***	***0.007***	***0.00***
Multi-market	0.005	0.003	0.06
pre 2000	0.002	0.006	0.68
pre 2005	-0.002	0.004	0.60
pre 2007	-0.005	0.003	0.13
pre 2009	0.008	0.003	0.02

level of returns is far more important than failing to capture the direction of change. To obtain more insight into the reasons for these failures, we repeat the regression in equation (5.4), firstly with $|1 - \beta|$ (using the deviation of β from one as a measure of failure to capture directional change) and secondly with α (as a measure of the systematic failure to capture the level of returns) as dependent variables.

The results are shown in Table 5.6. Considering the left-hand panel first, only 20% of the variability of the deviations $|1 - \beta|$ are explained by the properties of the ETFs. Considering the significant coefficients, we see that inverse and/or leveraged ETFs tend to achieve β closer to one than simple single market equity ETFs. The negative coefficients associated with the start dates indicate that β is better captured by those ETFs with earlier start dates, but the effect is not significant for start dates earlier than 2000.

Considering the right-hand panel of Table 5.6, we see that 64% of the variation in α, the more important source of underperformance, is explained by the properties of the ETFs. Inverse and/or leveraged ETFs tend to fail to reproduce the level of the benchmark return. There is a slight difference due to the type of benchmark, currencies tend to do better than the rest, multi-market tend to do slightly worse than the rest. The timing factors show deterioration in reproducing the level of return for start dates before January 2009 and an improvement for start dates before January 2007. This indicates that ETFs tended to replicate the level of return less well during the 2007-8 financial crisis.

In summary, a non-negligible proportion of ETFs underperform their benchmark in terms of return. Although there is a slight difference due to the type of index tracked, underperformance tends to be concentrated in inverse and leveraged ETFs. This underperformance is mainly due to failure to capture the level of returns of the benchmark, this

Table 5.6: Two regressions to decompose the underperformance of an ETF in terms of the properties of the ETF

Analysis of $\|1-\beta\|$	$R^2 = 0.20$			**Analysis of α**	$R^2 = 0.64$		
	Coefficient	*Standard Error*	*p-value*		*Coefficient*	*Standard Error*	*p-value*
Intercept	***0.322***	***0.038***	***0.00***	Intercept	0.008	0.005	0.06
$\boldsymbol{L_{-2}}$	***-0.191***	***0.039***	***0.00***	$\boldsymbol{L_{-2}}$	***-0.154***	***0.005***	***0.00***
$\boldsymbol{L_{-1}}$	***-0.254***	***0.071***	***0.00***	$\boldsymbol{L_{-1}}$	***-0.052***	***0.008***	***0.00***
$\boldsymbol{L_2}$	***-0.209***	***0.040***	***0.00***	$\boldsymbol{L_2}$	***-0.050***	***0.005***	***0.00***
ln(MV)	***0.035***	***0.005***	***0.00***	ln(MV)	0.000	0.001	0.45
Bond	***-0.106***	***0.040***	***0.01***	Bond	-0.003	0.005	0.51
Commodity	-0.066	0.040	0.10	Commodity	0.004	0.005	0.44
Currency	-0.053	0.062	0.40	***Currency***	***0.033***	***0.007***	***0.00***
Multi-market	-0.036	0.024	0.13	***Multi-market***	***-0.007***	***0.003***	***0.01***
pre 2000	0.004	0.053	0.94	pre 2000	-0.003	0.006	0.61
pre 2005	***-0.152***	***0.031***	***0.00***	pre 2005	0.002	0.004	0.57
pre 2007	***-0.114***	***0.028***	***0.00***	***pre 2007***	***0.011***	***0.003***	***0.00***
pre 2009	***-0.138***	***0.030***	***0.00***	***pre 2009***	***-0.021***	***0.004***	***0.00***

failure tended to be more pronounced during the 2007-8 financial crisis.

5.4.2 Underperformance in volatility

In Figure 5.6, we compare the volatility (% p.a.) of each ETF with that of its benchmark using all the data available for each benchmark. The diagonal line in Figure 5.6 divides the plot into two triangles. Points in the top left (upper) triangle of the plot show ETFs which are more volatile than their benchmarks ($\sigma_r > \sigma_B$); there are several ETFs where the excess volatility is very large. Points in the lower right triangle indicate instances where the ETF is less volatile than its benchmark; here the differences are small compared to the upper triangle. The distribution of the difference between the variance of the benchmark and the variance of the ETF, $(\sigma_B^2 - \sigma_r^2)$, is negatively skewed. Describing this distribution in more familiar volatility per annum; the lower quartile is equivalent to a volatility difference of 16.1% p.a. and the median is equivalent to a volatility difference of 4.8% p.a. (where the ETF volatility exceeds that of the benchmark); the upper quartile is equivalent to a volatility difference of 5.1% p.a. (where ETF volatility is less than that of the benchmark). Looking at these data in another way, 35% (51%) of ETFs yield an annual volatility within 1% p.a. (2% p.a.) of the volatility of their benchmark.

In order to try to explain why some ETFs fail to reproduce the volatility of the benchmarks, the following regression was estimated.

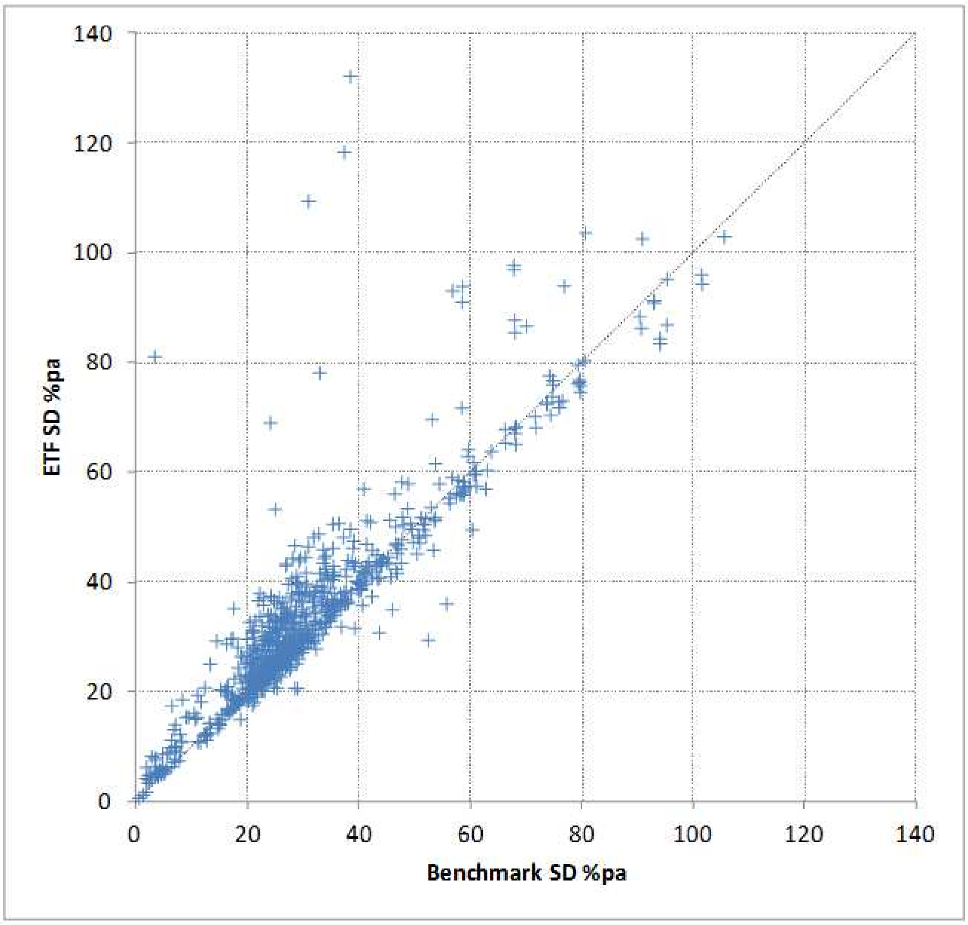

Figure 5.6: A comparison of ETF volatility with that of its benchmark. Volatility is calculated over all available data

$$\begin{aligned}\ln(\sigma_B^2/\sigma_r^2) = \quad & \theta_0 + \theta_1 L_{-2} + \theta_2 L_{-1} + \theta_3 L_2 + \theta_4 \ln(\text{Market Value}) \\ & +\theta_5 \text{Cat(Bond)} + \theta_6 \text{Cat(Commodity)} + \theta_7 \text{Cat(Currency)} \\ & +\theta_8 \text{Cat(Multi-market)} + \theta_9 (\text{pre } 2000) + \theta_{10} (\text{pre } 2005) \\ & +\theta_{11} (\text{pre } 2007) + \theta_{12} (\text{pre } 2009) + \text{error}\end{aligned} \tag{5.5}$$

where the indicators/variables are as previously defined above. The estimation results are summarised in Table 5.7. Note here that given regression equation (5.5) a negative coefficient (θ_1 to θ_{12}) indicates a factor that contributes to greater excess volatility, a positive coefficient a factor that reduces excess volatility.

The R^2 of the regression is only 7% so much of the variation is unexplained. However, the variables that have a significant negative effect (leading to greater excess volatility) are those that identify the category of multi-market tracker and inverse leveraged trackers.

From equation (5.3), we see that the difference in variance has two components: $(1 - \beta^2)\sigma_B^2$ due to failure to capture changes in benchmark returns; σ_ε^2 due to noise in the

Table 5.7: Regression to explain the difference between the volatility of an ETF and that of its benchmark in terms of the properties of the ETF

Analysis of $\ln(\sigma_B^2/\sigma_r^2)$ $R^2 = 0.07$			
	Coefficient	*Standard Error*	*p-value*
Intercept	-0.926	0.536	0.08
L_{-2}	***-2.900***	***0.554***	***0.00***
L_{-1}	0.447	0.997	0.65
L_2	0.107	0.563	0.85
ln(MV)	-0.041	0.070	0.56
Bond	-0.838	0.556	0.13
Commodity	-0.265	0.557	0.63
Currency	0.442	0.877	0.61
Multi-market	***-1.349***	***0.334***	***0.00***
pre 2000	-0.529	0.737	0.47
pre 2005	0.169	0.439	0.70
pre 2007	0.185	0.395	0.64
pre 2009	0.911	0.423	0.03

tracking process. These components are similar in importance, the median values are 0.7 and 0.9 (% per day)^2 respectively. We repeat the regression shown in equation (5.4), firstly with $(1-\beta^2)$ as the dependent variable and secondly with $\ln(\sigma_\varepsilon^2)$ as the dependent variable. These results are summarised in Table 5.8.

Table 5.8: Regressions looking at a decomposition of the difference between the volatility of an ETF and that of its benchmark in terms of the properties of the ETF

Analysis of $(1-\beta^2)$ $R^2 = 0.15$				**Analysis of $\ln(\sigma_\varepsilon^2)$** $R^2 = 0.32$			
	Coefficient	*Standard Error*	*p-value*		*Coefficient*	*Standard Error*	*p-value*
Intercept	***0.415***	***0.045***	***0.00***	Intercept	-0.041	0.172	0.81
L_{-2}	***-0.204***	***0.047***	***0.00***	***L_{-2}***	***0.496***	***0.178***	***0.01***
L_{-1}	***-0.249***	***0.084***	***0.00***	***L_{-1}***	***-0.804***	***0.32***	***0.01***
L_2	***-0.221***	***0.047***	***0.00***	L_2	-0.192	0.180	0.29
ln(MV)	***0.031***	***0.006***	***0.00***	ln(MV)	-0.020	0.022	0.36
Bond	-0.033	0.047	0.48	***Bond***	***-1.687***	***0.178***	***0.00***
Commodity	0.010	0.047	0.83	Commodity	0.431	0.178	0.02
Currency	0.016	0.074	0.83	***Currency***	***-1.088***	***0.281***	***0.00***
Multi-market	***-0.079***	***0.028***	***0.01***	***Multi-market***	***0.676***	***0.107***	***0.00***
pre 2000	0.050	0.062	0.42	***pre 2000***	***0.902***	***0.236***	***0.00***
pre 2005	***-0.118***	***0.037***	***0.00***	***pre 2005***	***-0.789***	***0.141***	***0.00***
pre 2007	***-0.132***	***0.033***	***0.00***	***pre 2007***	***-0.662***	***0.127***	***0.00***
pre 2009	***-0.147***	***0.036***	***0.00***	pre 2009	0.306	0.136	0.02

Considering the left-hand panel of Table 5.8, the departures of β^2 from unity are not well explained by the regression with a R^2 value of only 15%; there is a large amount of unexplained variation. This is a similar analysis to the left-hand panel of Table 5.6 with a different way of representing the departure of β from one; consequently the findings are similar with the extra suggestion that multi-market ETFs capture β better than single market ETFs.

Considering the right-hand panel of Table 5.8, where $\ln(\sigma_\varepsilon^2)$ is the dependent variable, we are seeking to explain the extent of the variability in an ETF's tracking of its benchmark. This regression has a R^2 value of 32%. This variability increases by the largest amount for multi-market ETFs. The variability decreases for bond and currency ETFs. The effect of inverse leverage (L_{-1} and L_{-2}) is mixed. The values for the start date coefficients suggest that variability increased during the period of the 2007-8 financial crisis.

To summarise, the volatility of the benchmark is exceeded by most ETFs; the discrepancy in volatility is caused in roughly equal proportions by failure to capture the direction and size of changes in returns of the benchmark (β) and by the variability in the tracking process (σ_ε^2). The type of benchmark mainly affected the variability, whereas the nature of the ETF (inverse and/or leveraged) mainly affected the capture of β. The variability in the tracking process increased during the recent financial crisis.

5.5 Summary and Conclusions

We have described the current composition of the market for ETFs and its rapid growth. The market value of the ETF market was estimated to exceed US$3 trillion (September 2011) having grown from US$2 billion in 2001. Equities represent 70% of this market value; single market equity ETFs are concentrated in US, China and Japan; the largest proportion of multi-market equity ETFs follow emerging markets. Commodities, mainly gold and oil, represent 20% of ETF market value. Approximately one in eight ETFs is either an inverse tracker, a leveraged tracker or both.

The availability of data limited the extent of our analysis. We were able to get a useful data history for 822 ETFs out of the 7198 ETFs active in September 2011. The accuracy with which ETFs replicate the behaviour of their benchmark is a mixed story. Using the data available to us from 1993 onwards, only 19% (29%) of ETFs reproduce both the mean return and the volatility of their benchmark within 1% p.a. (2% p.a.). Tracking accuracy tended to deteriorate during the 2007-8 financial crisis. We found that discrepancies in replicating the mean return of the benchmark tended to be associated with either leveraged or inverse (or both) ETFs. There was some evidence that discrepancies in replication of benchmark volatility was associated with multi-market ETFs; in contrast to bond and currency ETFs which tended to reproduce benchmark volatility more accurately than single market ETFs.

We have established that for many ETFs the replication of their benchmarks is imperfect. This means that in practice, if an ETF is used to hedge exposure to a market or a commodity under the assumption that the ETF will replicate its benchmark, then the discrepancies we have discussed mean that the hedge will be defective. To discover the reasons underlying these imperfections further work is needed.

One area of fruitful investigation would be to explore the mechanisms used to produce leveraged and inverse (synthetic) ETFs; identification of less successful mechanisms would improve overall benchmark tracking accuracy. A second area that could be explored would be the feasibility of providing regular publication of an up-to-date summary of the benchmark tracking accuracy of each ETF by an appropriate regulatory body, as this would very likely lead to improved tracking accuracy. Both of these areas would go some way to address the commonly articulated concerns of regulators with regard to complexity and opacity, especially with regard to synthetic ETFs.

Chapter 6

An optimisation approach to constructing an exchange-traded fund

6.1 Introduction

In Chapter 5, we presented an extensive survey and performance analysis of the ETF market up to 2011. In our study we observed constant underperformance for leveraged and inverse ETFs when compared to their benchmark, this is due especially to the nature of how they achieve the aimed multiple of the returns. ETF providers typically employ swaps/futures/derivative contracts that need to be rebalanced daily, incurring high transaction costs which undermine ETF performance.

For leveraged and inverse ETFs (so any ETF where $L \neq 1$) deciding the composition of the underlying portfolio to hold so as to (if possible) achieve the target return is a genuinely difficult task. For example, what portfolio of assets should one hold for a 2× ETF aiming to achieve twice the return on the S&P 500? In this chapter we consider the problem of deciding the portfolio of assets that should underlie an ETF. Note here that since index tracking (so $L = 1$) has been extensively considered in the literature (e.g. see Chen & Kwon (2012); Chiam et al. (2013); Garcia et al. (2011); Guastaroba & Speranza (2012); Mezali & Beasley (2013); Scozzari et al. (2013); Wang et al. (2012) for recent work) we primarily focus here only on cases where $L \neq 1$, so portfolios of assets aiming to give a return different from that of the benchmark index.

We formulate a mixed-integer nonlinear program for the problem of selecting the portfolio of assets underlying an ETF. Nonlinear formulations are often computationally challenging; in order to find (locally) optimal solutions we used the Minotaur solver

(Leyffer et al. (2013)), a toolkit for solving mixed-integer nonlinear optimisation problems. Our formulation allows both long and short positions.

The structure of this chapter is as follows. In Section 6.2 we give our formulation of the problem of deciding the portfolio of assets that should underlie an ETF. In Section 6.3 we give computational results for constructing ETFs for ten different problem instances derived from universes defined by S&P international equity indices, involving up to 500 assets. In Section 6.4, we compare our model with an alternative model for index-tracking. Finally in Section 6.5 we present our conclusions.

We would note here that, to the best of our knowledge, this chapter is the first in the literature to present a model for deciding the underlying assets to be held in order to construct an ETF which achieves a given multiple of benchmark return.

6.2 Problem formulation

In this section we formulate the problem of deciding the portfolio of assets that should underlie an ETF as a mixed-integer nonlinear program. We consider ETFs which have positive leverage with respect to their benchmark index and ETFs which have negative leverage (inverse, short, ETFs). Our formulation is a flexible one that incorporates decisions as to both long and short positions in assets, as well as including rebalancing and transaction cost.

6.2.1 Notation

We observe over time $0, 1, 2, \ldots, T$ the value of N assets. We are interested in selecting, at time T, the best set of K assets to hold (where $K \leq N$), as well as their appropriate quantities, to create an ETF whose return is a known multiple L of the benchmark index return. Here L is the ETF leverage factor, e.g. $L = 2$ for an ETF that aims to achieve twice the return of the benchmark index, $L = -1$ for an ETF that aims to achieve the negative of that return. Let:

V_{it}	be the value (price) of one unit of asset i at time t
I_t	be the value of the benchmark market index at time t
R_t	be the single period continuous time return for the index at time t, i.e. $R_t = \log_e(I_t/I_{t-1})$
X_i^L, X_i^S	be the number of units (≥ 0) of asset i held in the current ETF at time T in long/short positions respectively
C_{cash}	be the cash change in the ETF at time T ($C_{cash} > 0$ represents new cash to be invested in the ETF, $C_{cash} < 0$ represents cash to be taken out of the ETF)
C	be the total available (≥ 0) at time T to invest in the ETF, so $C = C_{cash} + \sum_{i=1}^{N} X_i^L V_{iT} - \sum_{i=1}^{N} X_i^S V_{iT}$
C^L, C^S	be the limits (> 0) on the total invested in long/short positions at time T
ε_i^L, ε_i^S	be the lower limits ($0 \leq \varepsilon_i^L, \varepsilon_i^S \leq 1$) on the proportion of C^L and C^S respectively invested in long/short positions in asset i if any position is taken
δ_i^L, δ_i^S	be the upper limits ($0 \leq \delta_i^L, \delta_i^S \leq 1$) on the proportion of C^L and C^S respectively invested in long/short positions in asset i
$F_i^L(\zeta,\theta), F_i^S(\zeta,\theta)$	be transaction cost functions that give the transaction cost (≥ 0) incurred for asset i in moving at time T from a position (long/short respectively) involving ζ units of the asset to θ units of the asset, where $F_i^L(\zeta,\theta) = F_i^S(\zeta,\theta) = 0$ if $\zeta = \theta$
γ	be the limit on the proportion of C that can be consumed by transaction cost (where $0 \leq \gamma \leq 1$)

Then our decision variables are:

x_i^L, x_i^S	the number of units (≥ 0) of asset i that we choose to hold in long/short positions respectively
C_t	the value of the ETF at time $t = 0, \ldots, T$
z_i^L, z_i^S	$\begin{cases} 1 & \text{if any of asset } i \text{ is held in long/short positions in the ETF} \\ 0 & \text{otherwise} \end{cases}$

Without significant loss of generality (since the sums of money involved are large) we allow $[x_i^L]$, $[x_i^S]$ to take fractional values.

6.2.2 Constraints

The constraints associated with our formulation of the problem are:

$$\varepsilon_i^L z_i^L \leq x_i^L V_{iT}/C^L \leq \delta_i^L z_i^L, \quad i = 1, \dots, N \tag{6.1}$$

$$\varepsilon_i^S z_i^S \leq x_i^S V_{iT}/C^S \leq \delta_i^S z_i^S, \quad i = 1, \dots, N \tag{6.2}$$

$$z_i^L + z_i^S \leq 1, \quad i = 1, \dots, N \tag{6.3}$$

$$\sum_{i=1}^{N} (z_i^L + z_i^S) = K \tag{6.4}$$

$$\sum_{i=1}^{N} x_i^L V_{iT} \leq C^L \tag{6.5}$$

$$\sum_{i=1}^{N} x_i^S V_{iT} \leq C^S \tag{6.6}$$

$$\sum_{i=1}^{N} F_i^L(X_i^L, x_i^L) + \sum_{i=1}^{N} F_i^S(X_i^S, x_i^S) \leq \gamma C \tag{6.7}$$

$$\sum_{i=1}^{N} x_i^L V_{iT} + \sum_{i=1}^{N} F_i^L(X_i^L, x_i^L) + \sum_{i=1}^{N} F_i^S(X_i^S, x_i^S) = C + \sum_{i=1}^{N} x_i^S V_{iT} \tag{6.8}$$

$$C_t = \sum_{i=1}^{N} x_i^L V_{it} - \sum_{i=1}^{N} x_i^S V_{it}, \quad t = 0, \dots, T \tag{6.9}$$

Equations (6.1) and (6.2) ensure that if an asset i is not held in a long/short position in the ETF ($z_i^L = 0, z_i^S = 0$) then x_i^L, x_i^S are also zero. If asset i is held then these equations ensure that the proportions of C^L, C^S held respect the limits defined. Equation (6.3) says that we cannot hold a long and short position in the same asset simultaneously. Equation (6.4) ensures that there are exactly K assets in the ETF. Equations (6.5) and (6.6) ensure that we respect the limits on the total invested in long/short positions. Equation (6.7) limits total transaction cost appropriately.

Equation (6.8) is the monetary balance equation and says that at time T we first generate cash $\sum_{i=1}^{N} x_i^S V_{iT}$ by shorting, add that to C and then use that sum to purchase long positions costing $\sum_{i=1}^{N} x_i^L V_{iT}$, as well as pay the total transaction cost, $\sum_{i=1}^{N} F_i^L(X_i^L, x_i^L) + \sum_{i=1}^{N} F_i^S(X_i^S, x_i^S)$. Note here that amending the constraints we have given above so as to have long-only ETFs without shorting is trivial (simply remove the variables associated with shorting).

In Equation (6.9) we have that at time t the ETF has x_i^L, $i = 1, \ldots, N$ in long positions, collectively worth $\sum_{i=1}^{N} x_i^L V_{it}$. The short positions x_i^S, $i = 1, \ldots, N$ represent obligations which have to be repaid (since in shorting, short selling, we borrow assets and have to return them). Collectively the short positions represent a (monetary) repayment of $\sum_{i=1}^{N} x_i^S V_{it}$ and so the value of the ETF at time t is as given in Equation (6.9). On a technical issue here note that Equation (6.9) is a valuation of the ETF portfolio at time t. The monetary value we would achieve at time t were we to liquidate the portfolio would potentially be less than this valuation due to transaction cost.

6.2.3 Objective function

In our formulation our objective is to minimise the tracking error (henceforth ***TE***), the averaged sum of the squared differences between ETF return and the required return, i.e.:

$$\text{minimise} \quad \text{TE} = \sum_{t=1}^{T} (\log_e(C_t/C_{t-1}) - LR_t)^2/T \tag{6.10}$$

It is clear from Equation (6.9) that the presence of short selling may mean that the value of the ETF becomes negative (or zero). In either case this would mean that the corresponding return in our objective function (Equation 6.10) is not defined. Numerically this would result in the optimisation software we are using giving an error. For this reason we impose the constraint that C_t, $t = 0, \ldots, T$ is greater than or equal to some small positive value. In the computational results reported later, since C is large, we (arbitrarily) imposed the constraint that $C_t \geq 10$, $t = 0, \ldots, T$.

Our formulation, minimise Equation (6.10) subject to Equations (6.1)-(6.9), is a mixed-integer program with a nonlinear objective function, so a MINLP (mixed-integer nonlinear program). Depending upon the nature of the transaction cost functions ($F_i^L(\zeta, \theta)$, $F_i^S(\zeta, \theta)$) which appear in Equations (6.7) and (6.8) the constraints may be linear or non-linear. In the simplest case when trading is free (so transaction cost is zero) the constraints are linear.

MINLPs are by their nature algorithmically and computationally challenging, since they combine discrete and continuous variables and are nonlinear. However packages for such programs have been improving (e.g. see Bussieck & Vigerske (2011)) and certainly our computational experience with Minotaur (Leyffer et al. (2013)) is that it is capable of solving quite large problems. With reference to our choice of the Minotaur solver please refer to Section 4.3.1.

6.2.4 Long/short fix

Similarly to Section 4.2.5, the amount that can be held in long/short positions is defined by the following equations:

$$C^S = \alpha C \tag{6.11}$$

$$C^L = (1 + \alpha)C \tag{6.12}$$

where α (where $0 \leq \alpha \leq 1$) is the proportion of C that can be held in short positions.

6.3 Computational results

In this section we present computational results for our approach to constructing an ETF. We used an Intel Xeon CPU E5-2640 @ 2.50GHz with 64GB of RAM with Linux as the operating system. The code was written in C++ and Minotaur 0.1.1 (Leyffer et al. (2013)) was used as the mixed-integer nonlinear solver in order to find (locally) optimal solutions to our formulation.

6.3.1 Data and methodology

In our computational experimentation we adopted successive periodic rebalancing over time and we used the same real-world historical weekly data as described in Section 3.4.1, the reader may refer to that section for details on the methodology. Two of the indices shown in Figure 3.1 are not used in this chapter. The first is the S&P World ex-US 700 index, since it was only assembled in 2004 (so we had limited data), and the second is the S&P Global 1200. Although we have data for this index for all years, the number of assets involved ($N = 1200$) meant that it was beyond the computational range of Minotaur (which consistently failed to find feasible solutions, even when given a substantial time limit).

With regard to parameter values we, unless otherwise stated, set the in-sample period $h = 52$ and the out-of-sample period $H = 13$. For the main results we used $\alpha = 0.50$, although in later sections we examine $\alpha = [0, 0.30, 0.50]$, which represent Long Only, 130/30 and 150/50 ETFs.

Unless otherwise stated we used $K = N$ as the number of assets to be in the ETF; $\varepsilon_i^L = \varepsilon_i^S = 0$ and $\delta_i^L = \delta_i^S = 1$ for $i = 1, \ldots, N$ as the proportion limits for each asset in the ETF. We set the initial ETF portfolio to consist of an investment of US$1 million in the first K assets held in equal proportions, with no short holdings. In other words $X_i^L = (1000000/K)/V_{i0}$ $i = 1, \ldots, K$; $X_i^L = 0$ $i = (K+1), \ldots, N$; $X_i^S = 0$ $i = 1, \ldots, N$.

For the main results we used $F_i^L(\zeta, \theta) = F_i^S(\zeta, \theta) = 0$ so transaction cost was zero. With these values the solver has complete freedom to choose the best (lowest tracking error) in-sample ETF amongst all possible solutions. However we do also present results below for the case where transaction cost was non-zero.

We imposed a time limit of $\max[2N, 900]$ seconds for each rebalance of our ETF, if the time limit is reached before Minotaur terminates we retrieve the best feasible solution found so far. We examined $L = -1$ and $L = 2$, which represent inverse and $2\times$ leveraged ETFs.

6.3.2 Results, inverse ETFs

Table 6.1 shows the tracking errors for $L = -1$ (inverse ETFs). The table presents in-sample and out-of-sample tracking errors. Given the total number of time intervals (400 weeks) and our choice of $H = 13$, for each instance there were a total of $\lfloor (400-h)/H+1 \rfloor = 27$ different rebalances, each with its own tracking error calculated from 52 different returns. The column labelled ***Num*** shows the number of successful rebalances (out of 27), meaning that Minotaur was able to find a feasible solution. Notice that this number deteriorates as the instance size grows. If a rebalance is unsuccessful we leave the ETF unchanged (i.e. the current ETF is carried forward).

There are four columns under ***In-sample TE***. The first column contains the ***Average*** tracking error over all successful rebalances, while columns ***Max*** and ***Min*** show the maximum (worst) and minimum (best) TE respectively. The ***SD*** (standard deviation) completes the remaining in-sample columns. Column ***Out-of-sample TE*** gives the total out-of-sample tracking error as calculated from the single time series of out-of-sample returns, while column ***Time*** shows the average CPU time (in seconds) needed to solve each rebalance. The average time here is calculated over successful rebalances (unsuccessful rebalances requiring $\max[2N, 900]$ seconds).

Our model tends to find better in-sample solutions for larger instances, even though they are harder (more time consuming) to solve. This can be seen when we examine the average in-sample tracking error, which shows a decreasing trend as the problem size increases. Larger instances offer a wider selection of assets to choose from, which compensates for the greater difficulty in finding a solution. The in-sample standard deviation also tends to be smaller for the larger instances.

For the eight instances where all rebalances were successful the out-of-sample tracking errors are all smaller than the average in-sample tracking error. For the two larger

Table 6.1: Tracking errors for the 150/50 case with $L = -1$

Index	Num	In-sample TE Average	Max	Min	SD	Out-of-sample TE	Time (seconds)
S&P Latin America 40	27	1.463×10^{-3}	6.926×10^{-3}	5.630×10^{-4}	1.261×10^{-3}	2.472×10^{-3}	0.65
S&P Asia 50	27	9.507×10^{-4}	7.504×10^{-3}	2.643×10^{-4}	1.392×10^{-3}	1.118×10^{-3}	7.01
S&P ASX 50	27	1.370×10^{-3}	9.653×10^{-3}	4.964×10^{-4}	1.768×10^{-3}	2.201×10^{-3}	0.52
S&P TSX 60	27	7.150×10^{-4}	1.337×10^{-3}	3.458×10^{-4}	2.323×10^{-4}	1.696×10^{-3}	5.11
S&P UK 125	27	2.702×10^{-4}	5.274×10^{-4}	5.295×10^{-5}	1.073×10^{-4}	1.489×10^{-3}	54.67
S&P Topix 150	27	1.013×10^{-3}	7.286×10^{-3}	2.279×10^{-4}	1.509×10^{-3}	2.314×10^{-3}	143.68
S&P Euro Zone 175	27	2.951×10^{-4}	5.216×10^{-4}	4.089×10^{-5}	1.081×10^{-4}	1.484×10^{-3}	52.99
S&P Euro Plus 225	27	2.712×10^{-4}	5.127×10^{-4}	2.415×10^{-5}	1.029×10^{-4}	1.430×10^{-3}	180.25
S&P Europe 350	23	1.201×10^{-4}	2.706×10^{-4}	1.628×10^{-8}	7.619×10^{-5}	1.311×10^{-3}	88.27
S&P US 500	19	4.197×10^{-5}	2.307×10^{-4}	4.812×10^{-7}	5.763×10^{-5}	1.785×10^{-3}	344.27
Average	25.8	6.511×10^{-4}	3.477×10^{-3}	2.016×10^{-4}	6.614×10^{-4}	1.730×10^{-3}	87.74

instances where we had some unsuccessful rebalances (and so carried the ETF forward without change) the out-of-sample tracking errors are larger than the average in-sample tracking errors.

6.3.3 Results, leveraged ETFs

In this section we examine the results for the case when $L = 2$, so leveraged ETFs. Table 6.2 summarises the results. Similarly to the results shown in Table 6.1, we can see that lower in-sample tracking errors are achieved for larger instances.

6.3.4 Varying h and α

In the previous results we used $h = 52$ and $\alpha = 0.5$. In this section we show what happens when we vary these parameters.

Table 6.3 shows average tracking error results over all instances for $\alpha = 0.5$ with $L = -1$ and $L = 2$, whilst varying h between 13 and 52 weeks. It should be noted that, under column ***Num***, the total number of rebalances for $h = [13, 26, 39, 52]$ is $[30, 29, 28, 27]$ respectively. The rows in this table for $h = 52$ are as the final rows in Tables 6.1 and 6.2, but are repeated here for ease of comparison.

In Table 6.3 we can observe that (both for $L = -1$ and $L = 2$) the higher the value of h, the better our in-sample and out-of-sample results are, both in terms of average tracking error and in terms of standard deviation in tracking error. As might be expected, the

Table 6.2: Tracking errors for the 150/50 case with $L = 2$

Index		In-sample TE				Out-of-sample	Time
	Num	Average	Max	Min	SD	TE	(seconds)
S&P Latin America 40	27	5.089×10^{-4}	9.067×10^{-3}	8.725×10^{-5}	1.712×10^{-3}	8.066×10^{-4}	0.65
S&P Asia 50	27	1.095×10^{-4}	2.270×10^{-4}	3.830×10^{-5}	5.990×10^{-5}	3.451×10^{-4}	1.17
S&P ASX 50	27	3.927×10^{-4}	3.337×10^{-3}	1.265×10^{-4}	6.019×10^{-4}	8.274×10^{-4}	8.31
S&P TSX 60	27	1.328×10^{-4}	2.788×10^{-4}	6.200×10^{-5}	5.299×10^{-5}	6.224×10^{-4}	10.80
S&P UK 125	27	4.890×10^{-5}	1.155×10^{-4}	1.042×10^{-5}	2.698×10^{-5}	5.719×10^{-4}	137.61
S&P Topix 150	26	2.939×10^{-4}	5.839×10^{-3}	1.966×10^{-6}	1.132×10^{-3}	7.809×10^{-4}	156.82
S&P Euro Zone 175	27	4.117×10^{-5}	9.930×10^{-5}	2.590×10^{-6}	2.843×10^{-5}	6.079×10^{-4}	27.85
S&P Euro Plus 225	27	2.388×10^{-5}	6.815×10^{-5}	2.075×10^{-6}	1.491×10^{-5}	5.698×10^{-4}	202.53
S&P Europe 350	27	1.018×10^{-6}	5.149×10^{-6}	1.502×10^{-19}	1.534×10^{-6}	4.215×10^{-4}	328.06
S&P US 500	15	1.334×10^{-7}	1.949×10^{-6}	3.354×10^{-17}	5.023×10^{-7}	5.963×10^{-4}	568.94
Average	25.7	1.553×10^{-4}	1.904×10^{-3}	3.311×10^{-5}	3.631×10^{-4}	6.150×10^{-4}	144.27

average solution time also increases as h increases. Comparing the results for inverse ETFs and leveraged ETFs we can observe that, for all values of h considered, the leveraged ETF gives a lower average tracking error (both in-sample and out-of-sample) than the inverse ETF.

Table 6.3: Average values for all instances with $\alpha = 0.5$

L	h		In-sample TE				Out-of-sample	Time
		Num	Average	Max	Min	SD	TE	(seconds)
$L = -1$	$h = 13$	30.0	3.204×10^{-3}	1.974×10^{-2}	2.253×10^{-5}	4.909×10^{-3}	4.030×10^{-3}	23.65
	$h = 26$	29.0	1.095×10^{-3}	7.878×10^{-3}	7.108×10^{-5}	1.820×10^{-3}	2.421×10^{-3}	59.57
	$h = 39$	27.9	7.380×10^{-4}	3.446×10^{-3}	1.429×10^{-4}	8.292×10^{-4}	2.131×10^{-3}	66.48
	$h = 52$	25.8	6.511×10^{-4}	3.477×10^{-3}	2.016×10^{-4}	6.614×10^{-4}	1.730×10^{-3}	87.74
$L = 2$	$h = 13$	29.9	3.776×10^{-3}	2.580×10^{-2}	2.931×10^{-10}	6.381×10^{-3}	2.659×10^{-3}	37.57
	$h = 26$	29.0	5.373×10^{-4}	3.260×10^{-3}	6.606×10^{-6}	9.482×10^{-4}	1.067×10^{-3}	85.56
	$h = 39$	28.0	3.577×10^{-4}	2.589×10^{-3}	2.052×10^{-5}	6.984×10^{-4}	8.039×10^{-4}	104.21
	$h = 52$	25.7	1.553×10^{-4}	1.904×10^{-3}	3.311×10^{-5}	3.631×10^{-4}	6.150×10^{-4}	144.27

We now fix $h = 52$ and vary α, the level of shorting allowed in the model. Table 6.4 shows results for both $L = -1$ and $L = 2$ where we examine $\alpha = [0.0, 0.3, 0.5]$, which represent Long Only, 130/30 and 150/50 ETFs. The rows in this table for 150/50 are as the final rows in Tables 6.1 and 6.2, but are repeated here for ease of comparison.

In Table 6.4 we can see that (both for $L = -1$ and $L = 2$) the higher the value of α, the better our in-sample and out-of-sample results are, both in terms of average tracking error and in terms of standard deviation in tracking error. Here the addition of shorting helps in reducing both in-sample and out-of-sample tracking errors, since we are adding the flexibility to short as compared with the Long Only case.

Comparing the results for inverse ETFs and leveraged ETFs we can see that, for all values of α considered, the leveraged ETF gives a lower average tracking error (both in-sample and out-of-sample), as well as a lower standard deviation in tracking error, as compared with the inverse ETF.

Table 6.4: Average values for all instances with $h = 52$

L	Case		In-sample TE				Out-of-sample	Time
		Num	Average	Max	Min	SD	TE	(seconds)
$L=-1$	Long Only	27.0	2.952×10^{-3}	8.082×10^{-3}	8.843×10^{-4}	2.002×10^{-3}	3.535×10^{-3}	0.11
	130/30	26.8	1.156×10^{-3}	5.120×10^{-3}	2.996×10^{-4}	1.206×10^{-3}	1.934×10^{-3}	42.87
	150/50	25.8	6.511×10^{-4}	3.477×10^{-3}	2.016×10^{-4}	6.614×10^{-4}	1.730×10^{-3}	87.74
$L=2$	Long Only	27.0	1.490×10^{-3}	5.486×10^{-3}	1.742×10^{-4}	1.418×10^{-3}	1.684×10^{-3}	0.14
	130/30	26.0	3.130×10^{-4}	2.654×10^{-3}	5.199×10^{-5}	5.989×10^{-4}	6.680×10^{-4}	89.87
	150/50	25.7	1.553×10^{-4}	1.904×10^{-3}	3.311×10^{-5}	3.631×10^{-4}	6.150×10^{-4}	144.27

6.3.5 Results with transaction cost

We now examine results for the case when we include transaction cost in the model. We used $F_i^L(X_i^L, x_i^L) = 0.01|X_i^L - x_i^L|V_{iT}$ and $F_i^S(X_i^S, x_i^S) = 0.01|X_i^S - x_i^S|V_{iT}$, i.e. transaction cost was one percent of the value of the assets traded (for both long and short trading). We used $\gamma = 0.01$ as the limit on the proportion of C that can be consumed by transaction cost, so one percent of the total cash available. For simplicity we just consider here results for 150/50 ETFs (so $\alpha = 0.50$, as in Table 6.1 and Table 6.2).

Table 6.5 and Table 6.6 show the results for inverse and leveraged ETFs. If we compare Table 6.5 to Table 6.1 and Table 6.6 to Table 6.2 we can see that there is a deterioration in both in-sample and out-of-sample tracking errors (as we would expect, since imposing a transaction cost constraint restricts the opportunities available at each rebalance). Computation times for these test problems do not seem to be significantly affected by the addition of the transaction cost constraint.

Comparing the average tracking errors for $L = -1$ with those for $L = 2$ in Tables 6.5 and 6.6 we can see that for each of the ten instances the average tracking error, both in-sample and out-of-sample, is larger for $L = -1$ than for $L = 2$.

Table 6.5: Tracking errors with transaction cost and $L = -1$

Index	Num	In-sample TE Average	Max	Min	SD	Out-of-sample TE	Time (seconds)
S&P Latin America 40	27	3.621×10^{-3}	9.701×10^{-3}	6.905×10^{-4}	2.360×10^{-3}	4.263×10^{-3}	0.37
S&P Asia 50	27	1.655×10^{-3}	3.811×10^{-3}	3.110×10^{-4}	1.012×10^{-3}	1.768×10^{-3}	1.03
S&P ASX 50	27	3.381×10^{-3}	9.079×10^{-3}	5.718×10^{-4}	2.416×10^{-3}	2.848×10^{-3}	1.03
S&P TSX 60	27	1.701×10^{-3}	8.974×10^{-3}	3.673×10^{-4}	2.096×10^{-3}	1.798×10^{-3}	34.79
S&P UK 125	27	9.329×10^{-4}	3.273×10^{-3}	1.772×10^{-4}	8.257×10^{-4}	1.763×10^{-3}	81.29
S&P Topix 150	27	1.828×10^{-3}	4.682×10^{-3}	2.998×10^{-4}	1.260×10^{-3}	2.921×10^{-3}	118.74
S&P Euro Zone 175	27	7.019×10^{-4}	2.110×10^{-3}	1.665×10^{-4}	4.698×10^{-4}	1.859×10^{-3}	144.59
S&P Euro Plus 225	25	6.766×10^{-4}	2.536×10^{-3}	1.051×10^{-4}	5.855×10^{-4}	1.588×10^{-3}	136.85
S&P Europe 350	24	3.084×10^{-4}	1.301×10^{-3}	1.678×10^{-5}	3.232×10^{-4}	1.340×10^{-3}	120.63
S&P US 500	22	4.587×10^{-4}	3.015×10^{-3}	2.921×10^{-5}	7.280×10^{-4}	1.641×10^{-3}	96.12
Average	26.0	1.526×10^{-3}	4.848×10^{-3}	2.735×10^{-4}	1.208×10^{-3}	2.179×10^{-3}	73.54

Table 6.6: Tracking errors with transaction cost and $L = 2$

Index	Num	In-sample TE Average	Max	Min	SD	Out-of-sample TE	Time (seconds)
S&P Latin America 40	27	8.489×10^{-4}	5.269×10^{-3}	1.218×10^{-4}	1.241×10^{-3}	9.839×10^{-4}	0.76
S&P Asia 50	27	1.425×10^{-3}	4.227×10^{-3}	5.258×10^{-5}	1.274×10^{-3}	1.051×10^{-3}	0.80
S&P ASX 50	27	1.389×10^{-3}	4.061×10^{-3}	2.086×10^{-4}	1.144×10^{-3}	9.921×10^{-4}	1.46
S&P TSX 60	27	1.636×10^{-4}	3.572×10^{-4}	6.413×10^{-5}	8.136×10^{-5}	6.147×10^{-4}	8.77
S&P UK 125	27	1.861×10^{-4}	1.337×10^{-3}	3.513×10^{-5}	2.953×10^{-4}	5.460×10^{-4}	46.05
S&P Topix 150	27	2.899×10^{-4}	1.170×10^{-3}	4.835×10^{-5}	3.329×10^{-4}	6.622×10^{-4}	189.51
S&P Euro Zone 175	27	1.120×10^{-4}	8.345×10^{-4}	1.407×10^{-5}	1.543×10^{-4}	5.750×10^{-4}	134.96
S&P Euro Plus 225	27	6.550×10^{-5}	1.809×10^{-4}	1.320×10^{-5}	4.765×10^{-5}	5.141×10^{-4}	148.44
S&P Europe 350	20	5.471×10^{-5}	5.763×10^{-4}	4.050×10^{-6}	1.246×10^{-4}	4.375×10^{-4}	284.91
S&P US 500	21	5.847×10^{-6}	7.031×10^{-5}	1.958×10^{-12}	1.520×10^{-5}	4.191×10^{-4}	687.77
Average	25.7	4.541×10^{-4}	1.808×10^{-3}	5.619×10^{-5}	4.710×10^{-4}	6.796×10^{-4}	150.34

6.3.6 Results with restrictions on asset holdings

In the results presented above we did not enforce minimum or maximum limits on the proportion of the ETF that could be held in any asset. In this section we investigate the effect of this. We used $\varepsilon_i^S = \varepsilon_i^L = 0.25/K$ and $\delta_i^S = \delta_i^L = 2/K$ and as well set $K = 0.2N$ to simultaneously restrict the number of assets in the ETF. As before we just consider here results for 150/50 ETFs (so $\alpha = 0.50$, as in Table 6.1 and Table 6.2).

Table 6.7 and Table 6.8 show the results for inverse and leveraged ETFs. If we compare Table 6.7 to Table 6.1 and Table 6.8 to Table 6.2 we can see that there is a deterioration in both in-sample and out-of-sample tracking errors. Computation times are significantly higher in Table 6.7 and Table 6.8 than in Table 6.1 and Table 6.2.

Table 6.7: Tracking errors with asset limits, $L = -1$

Index		In-sample TE				Out-of-sample	Time
	Num	Average	Max	Min	SD	TE	(seconds)
S&P Latin America 40	27	3.574×10^{-3}	9.570×10^{-3}	9.608×10^{-4}	2.648×10^{-3}	4.118×10^{-3}	186.63
S&P Asia 50	27	1.532×10^{-3}	3.141×10^{-3}	4.812×10^{-4}	7.968×10^{-4}	1.802×10^{-3}	371.29
S&P ASX 50	27	1.828×10^{-3}	5.551×10^{-3}	7.639×10^{-4}	1.187×10^{-3}	2.642×10^{-3}	402.95
S&P TSX 60	27	9.598×10^{-4}	1.807×10^{-3}	5.265×10^{-4}	3.273×10^{-4}	1.887×10^{-3}	256.64
S&P UK 125	27	9.577×10^{-4}	2.648×10^{-3}	3.216×10^{-4}	5.407×10^{-4}	2.151×10^{-3}	579.67
S&P Topix 150	27	1.745×10^{-3}	6.375×10^{-3}	4.060×10^{-4}	1.405×10^{-3}	3.044×10^{-3}	578.14
S&P Euro Zone 175	27	1.017×10^{-3}	2.463×10^{-3}	2.978×10^{-4}	6.089×10^{-4}	1.893×10^{-3}	584.87
S&P Euro Plus 225	27	8.916×10^{-4}	2.083×10^{-3}	2.830×10^{-4}	5.081×10^{-4}	1.866×10^{-3}	602.89
S&P Europe 350	27	6.302×10^{-4}	1.731×10^{-3}	2.125×10^{-4}	4.166×10^{-4}	1.672×10^{-3}	636.90
S&P US 500	27	6.405×10^{-4}	1.702×10^{-3}	1.039×10^{-4}	5.332×10^{-4}	1.608×10^{-3}	945.34
Average	27.0	1.378×10^{-3}	3.707×10^{-3}	4.357×10^{-4}	8.971×10^{-4}	2.268×10^{-3}	514.53

Table 6.8: Tracking errors with asset limits, $L = 2$

Index		In-sample TE				Out-of-sample	Time
	Num	Average	Max	Min	SD	TE	(seconds)
S&P Latin America 40	27	9.506×10^{-4}	3.471×10^{-3}	2.975×10^{-4}	8.073×10^{-4}	1.161×10^{-3}	205.63
S&P Asia 50	27	7.858×10^{-4}	2.783×10^{-3}	2.743×10^{-4}	6.548×10^{-4}	8.349×10^{-4}	307.56
S&P ASX 50	27	7.700×10^{-4}	3.487×10^{-3}	2.479×10^{-4}	6.648×10^{-4}	8.935×10^{-4}	306.06
S&P TSX 60	27	2.909×10^{-4}	7.609×10^{-4}	1.260×10^{-4}	1.602×10^{-4}	9.255×10^{-4}	242.43
S&P UK 125	27	2.692×10^{-4}	5.796×10^{-4}	4.048×10^{-5}	1.536×10^{-4}	6.665×10^{-4}	635.19
S&P Topix 150	27	2.521×10^{-4}	5.444×10^{-4}	8.831×10^{-5}	1.451×10^{-4}	7.013×10^{-4}	668.76
S&P Euro Zone 175	27	1.674×10^{-4}	3.995×10^{-4}	4.422×10^{-5}	8.503×10^{-5}	5.912×10^{-4}	833.71
S&P Euro Plus 225	27	1.724×10^{-4}	6.702×10^{-4}	3.083×10^{-5}	1.575×10^{-4}	5.615×10^{-4}	835.08
S&P Europe 350	27	1.873×10^{-4}	6.423×10^{-4}	3.131×10^{-5}	1.479×10^{-4}	4.913×10^{-4}	680.99
S&P US 500	26	1.644×10^{-4}	7.371×10^{-4}	1.598×10^{-5}	1.884×10^{-4}	4.697×10^{-4}	1088.60
Average	26.9	4.010×10^{-4}	1.407×10^{-3}	1.197×10^{-4}	3.165×10^{-4}	7.296×10^{-4}	580.40

Comparing the average tracking errors for $L = -1$ with those for $L = 2$ in Tables 6.7 and 6.8 we can see that for each of the ten instances the average tracking error, both in-sample and out-of-sample, is larger for $L = -1$ than for $L = 2$.

It is noticeable that 19 of the 20 instances in Table 6.7 and Table 6.8 have 27 rebalances (so all rebalances were successful), the remaining instance (the S&P US 500 with $L = 2$) has 26 rebalances. Having a value of 26 in the Num column for the S&P US 500 with $L = 2$ indicates that at just one rebalance Minotaur could not find a solution and so the current ETF had to be carried forward. This sharply contrasts with Table 6.1 and Table 6.2 and indicates that (albeit at the expense of computation time) the reduction in the search space brought about by the imposition of constraints on asset holdings and the number of assets held enables Minotaur to more successfully solve the problem.

6.4 Results, index tracking ($L = 1$)

Since index tracking problems have been widely studied in the literature, in this chapter we emphasised leveraged and inverse ETFs. In this section, however, we compare our MINLP model solved by Minotaur with $L = 1$ (so an index tracking optimisation model) to the Population Heuristic (PH) presented in Beasley et al. (2003).

Beasley et al. (2003) constructed five test data sets by considering the assets involved in five different capital market indices drawn from around the world. Specifically they considered the Hang Seng (Hong Kong), DAX 100 (Germany), FTSE 100 (UK), S&P 100 (USA) and Nikkei 225 (Japan). The data consists of weekly prices from March 1992 to September 1997 for the assets in these indices, obtained via Thomson Reuters Datastream (2013). Stocks with missing values were dropped. There were 291 values for each stock from which to calculate (weekly) returns.

For the comparisons presented below we used the same configuration as in Beasley et al. (2003). We start the optimisation process with an initial tracking portfolio of the first $K = 10$ assets in equal proportions, i.e. $X_i = (C/K)/V_{i0}, i = 1, \ldots, K$; $X_i = 0 \; \forall i > K$. We set $C = 10^6$ and the model as Long Only. The index and each stock (composed of 291 prices) have 290 returns, we set $h = H = 145$ so half the time period was considered in-sample, and the other half out-of-sample, with no rebalancing of the portfolio. Since the problems solved are large (having $h = 145$), we set a time limit of 4800 seconds.

6.4.1 Artificial index

The first comparison regards how close the results are to optimality. Their heuristic cannot guarantee that the problem will be solved optimally. Hence, in the absence of benchmark optimal solutions against which to compare heuristic solutions, they artificially set the index value as $I_t = \sum_{i=N-K+1}^{N} V_{it}$, so that the index value is composed of the last K

assets; and solved the problem in the situation corresponding to zero transaction cost, $\varepsilon_i = 0$ and $\delta_i = 1\ \forall i$. In this case they know that it is possible to track the index with a tracking error of zero.

Table 6.9 compares the tracking errors obtained by Minotaur with those of the PH. The average time needed by PH was, in 2003, approximately 400 seconds. Due to the difficult nature of MINLPs, in terms of execution time Minotaur is clearly expected to be at a disadvantage, and we do not include PH times in the table as it is not a fair comparison due to the year in which they were executed.

Table 6.9: Artificial index, Minotaur versus PH

Index	N	Tracking error		Time (seconds)
		Minotaur	PH	
Hang Seng	31	8.538×10^{-10}	2.560×10^{-8}	49.33
DAX	85	7.480×10^{-11}	6.405×10^{-5}	2457.47
FTSE	89	2.298×10^{-10}	1.911×10^{-4}	290.62
S&P	98	9.761×10^{-10}	1.528×10^{-4}	291.64
Nikkei	225	7.267×10^{-4}	2.294×10^{-4}	4800.52

In terms of tracking errors, for the first four instances, Minotaur found better solutions than PH and finished its execution within the preset time limit. The solutions obtained are very close to zero. For the Nikkei 225, Minotaur was not able to solve the problem before the time limit and the best solution found was slightly worse than the heuristic solution. Clearly for large problems here choosing a heuristic is advantageous over solving with MINLP.

6.4.2 Real index

In this section we compare, for different transaction cost limits, the PH and Minotaur results. We simulate the exact experiments as conducted in Beasley et al. (2003), so we set $\varepsilon_i = 0.01$ and $\delta_i = 1\ \forall i$. For transaction cost, we used $F_i^L(X_i^L, x_i^L) = 0.01|X_i^L - x_i^L|V_{iT}$ and $F_i^S(X_i^S, x_i^S) = 0.01|X_i^S - x_i^S|V_{iT}$.

Table 6.10 compares the Minotaur results to PH, both in-sample and out-of-sample. For each instance, we test $\gamma = [0, 0.0025, 0.0050, 0.0075, 0.0100]$. Setting $\gamma = 0$ essentially means that we cannot make any changes to our portfolio and hence the results are exactly the same as those obtained by PH. The higher γ is, the more freedom there is to choose a better tracking portfolio. On the other hand, the problem is less constrained and more difficult to solve.

For both in-sample and out-of-sample, there are two columns in the Table: ***T. error***, containing the tracking errors found by Minotaur, and ***% of PH***, meaning how Minotaur results compare (in terms of percentage) to PH results. For example, taking the Hang Seng index, $\gamma = 0.0025$, Minotaur found an in-sample tracking error that is 96.56% of the tracking error found by PH. We can verify in the Table that, with few exceptions, the higher γ is, the better our in-sample tracking errors are. In most cases this trend is also repeated out-of-sample. Computational times also grow as γ grows.

For all instances solved by Minotaur within the time limit, the in-sample tracking error was better than that of PH, with the exception of Nikkei 225, $\gamma = 0.0025$. Minotaur results were slightly better overall, with an average tracking error of 99.21% of the average PH tracking error. The out-of-sample results, however, are a mixed picture. A better in-sample result does not necessarily turn into a better out-of-sample tracking error. On average, Minotaur out-of-sample results were 108.09% worse than those obtained by PH. The long out-of-sample period (145 weeks), with no rebalance, can perhaps explain these mixed results.

In general, Minotaur is, in manageable times, able to compete with PH. However, for larger problems (large instances or long in-sample periods) we can benefit from using a heuristic method.

6.5 Conclusions

In this chapter we have considered the problem of deciding the portfolio of assets that should underlie an exchange-traded fund in order to achieve a given multiple L of the return on a benchmark index. We formulated this problem as a mixed-integer nonlinear program and our formulation took into account long/short positions, rebalancing and transaction cost.

Computational results, obtained using the Minotaur solver, were given for problems derived from universes defined by S&P international equity indices, involving up to 500 assets. These indicated that, for the instances we examined, inverse ($L = -1$) ETFs were associated with larger tracking errors than leveraged ($L = 2$) ETFs.

Computational results also indicated that, for the instances examined, we were able to deal with transaction cost and asset limits, albeit at an increase in computation time when asset limits were imposed.

To the best of our knowledge, this chapter is the first in the literature to present a model for deciding the underlying assets to be held in order to construct an exchange-traded fund which achieves a given multiple of benchmark return.

Table 6.10: In-sample and out-of-sample tracking errors, Minotaur versus PH

Index	N	T. cost	In-Sample		Out-of-Sample		Time (s)
		limit γ	T. Error	% of PH	T. Error	% of PH	
		0.0000	1.028×10^{-3}	100.00	1.267×10^{-3}	100.00	0.00
		0.0025	8.425×10^{-4}	96.56	1.060×10^{-3}	99.44	12.99
Hang Seng	31	0.0050	6.739×10^{-4}	98.54	8.783×10^{-4}	100.80	14.84
		0.0075	5.199×10^{-4}	99.05	6.453×10^{-4}	100.62	20.60
		0.0100	4.099×10^{-4}	89.32	4.300×10^{-4}	75.31	45.74
		0.0000	1.173×10^{-3}	100.00	2.049×10^{-3}	100.00	0.00
		0.0025	9.500×10^{-4}	99.16	1.904×10^{-3}	107.21	148.46
DAX	85	0.0050	7.426×10^{-4}	96.87	1.635×10^{-3}	101.36	4800.06
		0.0075	6.339×10^{-4}	106.73	1.352×10^{-3}	96.99	4800.03
		0.0100	4.345×10^{-4}	96.23	1.217×10^{-3}	106.20	4800.29
		0.0000	1.021×10^{-3}	100.00	9.584×10^{-4}	100.00	0.00
		0.0025	8.023×10^{-4}	95.70	9.357×10^{-4}	101.52	305.63
FTSE	89	0.0050	6.584×10^{-4}	99.70	8.390×10^{-4}	102.94	4800.27
		0.0075	5.475×10^{-4}	94.48	8.020×10^{-4}	104.02	4800.18
		0.0100	5.150×10^{-4}	99.23	8.121×10^{-4}	125.70	4800.18
		0.0000	1.038×10^{-3}	100.00	1.032×10^{-3}	100.00	0.00
		0.0025	7.656×10^{-4}	94.45	8.096×10^{-4}	114.89	52.96
S&P	98	0.0050	6.130×10^{-4}	95.93	7.651×10^{-4}	118.60	4800.04
		0.0075	5.146×10^{-4}	101.98	5.838×10^{-4}	87.82	4800.31
		0.0100	5.248×10^{-4}	108.92	7.470×10^{-4}	115.81	4800.99
		0.0000	7.803×10^{-4}	100.00	8.213×10^{-4}	100.00	0.00
		0.0025	6.017×10^{-4}	100.17	9.976×10^{-4}	100.30	3807.26
Nikkei	225	0.0050	5.547×10^{-4}	100.36	7.219×10^{-4}	102.41	4800.41
		0.0075	5.025×10^{-4}	100.00	7.767×10^{-4}	92.17	4800.91
		0.0100	4.895×10^{-4}	106.75	7.092×10^{-4}	108.41	4800.72
Average				99.21%		108.09%	

Chapter 7

Conclusions

7.1 Summary

The aim of this thesis was to examine three different portfolio optimisation problems that, albeit being relatively well known, are not clearly defined from a scientific point of view. The three problems are that of defining an Absolute Return Portfolio (ARP), a Market Neutral Portfolio (MNP) and the basket underlying an Exchange-Traded Fund (ETF). We have presented a mathematical model for each of these three problems, along with extensive computational results.

In Chapter 2 we gave a review of previous work on ARPs, MNPs and ETFs, as well as on portfolio optimisation theory in general. We observed that there is no single accepted mathematical description to any of these problems. Overall, we summarise the works on these three models as very fragmented, with different models and different data result in isolated papers, and great difficulty in connecting them in a mathematical/data sense. Many of the papers seen in the literature lack detailed computational results.

In Chapter 3 we have considered the problem of selecting an ARP. We presented a three-stage mixed-integer zero-one program for the problem that explicitly considers transaction costs associated with trading. We extended our approach to present models for enhanced indexation (relative return) portfolios and for portfolios that are a mix of absolute and relative return. Computational results were given for portfolios derived from universes defined by S&P international equity indices which indicated that all three models produced good quality results and that the computation time required was not significant.

In Chapter 4 we considered the problem of constructing a MNP where we can hold both long and short positions in assets. We formulated this problem as a mixed-integer nonlinear program, minimising the absolute value of the correlation between portfolio return and index return, and solved it using the Minotaur software package. Computational results were presented for the same set of instances as used in Chapter 3. These indicated that in-sample we could achieve very low correlations (in many cases zero correlation) in reasonable computation times. Out-of-sample correlations were higher, but for the majority of cases examined the market neutral portfolios constructed using the approach given out-performed their benchmark indices. Computational results, for the test problems considered, indicated that the model proposed out-performed an alternative approach based on minimising the absolute value of regression slope (the zero-beta approach). We also compared our approach with the performance of seven funds that adopt market neutral strategies with respect to the S&P 500, this comparison indicated that in general we had comparable to worse correlations but significant better performance.

In Chapter 5 we described the current composition of the market for ETFs and its rapid growth. The market value of the ETF market was estimated to exceed US$3 trillion (September 2011). Equities represent 70% of this market value; commodities, mainly gold and oil, represent 20% of ETF market value. Approximately one in eight ETFs is either an inverse tracker, a leveraged tracker or both.

The availability of data limited the extent of our analysis. We were able to get a useful data history for 822 ETFs out of the 7198 ETFs active in September 2011. The accuracy with which ETFs replicate the behaviour of their benchmark is a mixed story. Using the data available to us from 1993 onwards, only 19% (29%) of ETFs reproduce both the mean return and the volatility of their benchmark within 1% p.a. (2% p.a.). Tracking accuracy tended to deteriorate during the 2007-8 financial crisis. We found that discrepancies in replicating the mean return of the benchmark tended to be associated with either leveraged or inverse (or both) ETFs. We have established that for many ETFs the replication of their benchmarks is imperfect. This means that in practice, if an ETF is used to hedge exposure to a market or a commodity under the assumption that the ETF will replicate its benchmark, then the discrepancies we have highlighted mean that the hedge will be defective.

Following our conclusions in Chapter 5, in Chapter 6 we considered the problem of deciding the portfolio of assets that should underlie an exchange-traded fund in order to achieve a given multiple L of the return on a benchmark index. We formulated this problem as a mixed-integer nonlinear program and our formulation took into account long/short positions, rebalancing and transaction cost. Computational results, obtained using the Minotaur solver, were given for the instance set described in Chapter 3, but limited to up to 500 assets due to difficulties in solving larger instances of our model. These indicated that, for the instances we examined, inverse ($L = -1$) ETFs were associated with larger tracking errors than leveraged ($L = 2$) ETFs. Computational results also indicated that we were able to deal with transaction cost and asset limits, albeit at an increase in computation time when asset limits were imposed. We also compared our model with $L = 1$ to a population heuristic presented by Beasley et al. (2003).

7.2 Contribution to knowledge

Chapter 2 demonstrated that we have read, and are familiar with, the relevant scientific literature with regard to ARPs, MNPs and ETFs. All of the mathematical/computational work presented in Chapters 3, 4 and 6 is, to the best of our knowledge, an original contribution to knowledge.

In Chapter 3, we clearly defined an ARP via the three-stage objective function. In Chapter 4, we presented a nonlinear model that minimises the correlation between the portfolio chosen and the return of the benchmark index, and which we could solve with up to 1200 stocks. To the best of our knowledge this chapter is the first in the literature to directly address correlation minimisation. In Chapter 6, we presented what we believe to be the first model for deciding the underlying assets to be held in order to construct an ETF which achieves a given multiple of benchmark return.

Moreover, in Chapter 5, we performed the largest ETF snapshot/analysis presented to date in the literature, with 8192 ETFs with a total market value in September 2011 of US$2.96 trillion. We also present, for a subset of 822 ETFs, a regression based performance analysis to gain insight into the relationship between ETF characteristics and their performance in replicating both benchmark return and benchmark volatility. This analysis was based on over 1.1m daily return observations.

The work in each chapter presented in this thesis has been submitted as four different papers to appropriate journals.

7.3 Future research

There are a number of ways in which we can take our work further. One approach that was discussed in Section 3.4.8 but remains valid for all three models presented in this thesis is with respect to higher-frequency price data. We have used weekly asset price data which had been manually adjusted to account for changes in index composition, removing susceptibility to the influence of survivor bias. Whilst obtaining price data (say at daily frequency) from commercial databases is relatively easy there is significant effort involved in adjusting such data to account for index composition changes. This effort remains a topic for future research.

Another line of inquiry discussed in Section 3.4.8, but which may also be applicable to MNPs, is a rebalance/liquidate/reinvest trading strategy, in which we may choose to liquidate our position if it does not seem to be advantageous to hold long/short positions (i.e. simply hold cash). We also leave this possibility for future research.

Another possible line of research that is valid for all three models is to use some sort of scenario prediction instead of basing our decision only on the immediate past. There are several techniques to generate future scenarios for asset prices, such as Geometric Brownian Motion (Freedman (1972)) or GARCH models (Bollerslev (1986)). Whether or not our models could benefit from more sophisticated predictions of prices is also a topic for future research.

Regarding MNPs, we have not performed computational experiments when transaction costs are included, which is left as future research. We have also discussed in Section 4.4.2 the possibility of adopting other correlation measures as opposed to the Pearson product-moment coefficient. Due to the more complex nature of other proposed coefficients, we also leave the exploration of these alternative coefficients for future work.

Our ETF model is restricted to trading on assets. Leveraged and inverse ETFs typically employ swaps/futures/derivative contracts as means to achieve the multiple of their benchmark return. As future research, it may well be possible to expand our model to include these more complex financial products.

For our nonlinear models, we could explore other solvers or develop specialised exact algorithms, or heuristic methods, so as to reduce computational times. For the ETF basket model in Chapter 6, computational time was an issue since we could not find meaningful results for the S&P Global 1200 index. Also, we observed in Section 6.4 that for larger instances a heuristic method can be beneficial.

References

Aber, J. W., Li, D., & Can, L. (2009). Price volatility and tracking ability of ETFs. *Journal of Asset Management*, *10*(4), 210–221.

Ackert, L. F., & Yisong, S. T. (2008). Arbitrage, liquidity, and the valuation of exchange traded funds. *Financial Markets, Institutions & Instruments*, *17*(5), 331–362.

Agapova, A. (2011). Conventional mutual index funds versus exchange-traded funds. *Journal of Financial Markets*, *14*(2), 323–343.

Agrrawal, P., & Clark, J. M. (2007). ETF betas: A study of their estimation sensitivity to varying time intervals. *ETFs and Indexing*, *1*, 96–103.

Al-Sharkas, A. (2005). The return in hedge-fund strategies. *Journal of Business*, *10*(3), 217–231.

Alexander, C. (2001). *Market models: A guide to financial data analysis*. John Wiley & Sons.

Alexander, C., & Barbosa, A. (2008). Hedging index exchange traded funds. *Journal of Banking & Finance*, *32*(2), 326–337.

Alexander, G. J. (1977). An algorithmic approach to deriving the minimum-variance zero-beta portfolio. *Journal of Financial Economics*, *4*(2), 231–236.

Amenc, N., Goltz, F., & Grigoriu, A. (2008). Constructing absolute return funds with ETFs: A dynamic risk–budgeting approach. *ETFs and Indexing*, *1*, 37–46.

Amenc, N., Malaise, P., & Martellini, L. (2004). Revisiting core-satellite investing: A dynamic model of relative risk management. *Journal of Portfolio Management*, *31*(1), 64–75.

Avellaneda, M., & Lee, J. H. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, *10*(7), 761–782.

Avellaneda, M., & Zhang, S. (2010). Path-dependence of leveraged ETF returns. *SIAM Journal on Financial Mathematics*, *1*(1), 568–603.

Badrinath, S. G., & Gubellini, S. (2011). On the characteristics and performance of long-short, market-neutral and bear mutual funds. *Journal of Banking & Finance*, *35*(7), 1762–1776.

Baele, L., & Londono, J. M. (2013). Understanding industry betas. *Journal of Empirical Finance*, *22*, 30–51.

Bank of England Financial Stability Report (2011). June, Issue No. 29. Bank of England, London, UK. Available from http://www.bankofengland.co.uk/publications/Docu ments/fsr/2011/fsrfull1106.pdf, last accessed October 25th 2013.

Bank of England Financial Stability Report (2012). June, Issue No. 31. Bank of England, London, UK. Available from http://www.bankofengland.co.uk/publications/Docu ments/fsr/2012/fsrfull1206.pdf, last accessed October 25th 2013.

Baronyan, S. R., Boduroglu, I. I., & Sener, E. (2010). Investigation of stochastic pairs trading strategies under different volatility regimes. *The Manchester School*, *78*(1), 114–134.

Barro, D., & Canestrelli, E. (2010). Tracking error with minimum guarantee constraints. In Corazza, Marco and Pizzi, Claudio (Ed.) *Mathematical and statistical methods for actuarial sciences and finance*, (pp. 13–21). Springer Milan.

Beasley, J. E. (2013). Portfolio optimisation: models and solution approaches. In *Tutorials in Operations Research, Chapter 11*, vol. 10, (pp. 201–221). Published by Informs.

Beasley, J. E., Meade, N., & Chang, T. J. (2003). An evolutionary heuristic for the index tracking problem. *European Journal of Operational Research*, *148*(3), 621–643.

Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, *23*(4), 769–805.

Berkelaar, A., Dert, C., Oldenkamp, B., & Zhang, S. (2002). A primal dual decomposition-based interior point approach to two-stage stochastic linear programming. *Operations Research*, *50*(5), 904–915.

Berkelaar, A., Gromicho, J., Kouwenberg, R., & Zhang, S. (2005). A primal-dual decomposition algorithm for multistage stochastic convex programming. *Mathematical Programming Series A*, *104*(1), 153–177.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, *45*(3), 444–455.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, *48*(5), 28–43.

BlackRock (2011). ETP landscape, Global Handbook, Q4 2011. Available at http://www2.blackrock.com/content/groups/internationalsite/documents/literature/etfl_globalhandbook_q411_ca.pdf, last accessed October 25th 2013.

Blitz, D., & Huij, J. (2012). Evaluating the performance of global emerging markets equity exchange-traded funds. *Emerging Markets Review*, *13*(2), 149–158.

Blitz, D., Huij, J., & Swinkels, L. (2012). The performance of European index funds and exchange-traded funds. *European Financial Management*, *18*(4), 649–662.

Boehmer, B., & Boehmer, E. (2003). Trading your neighbor's ETFs: Competition or fragmentation? *Journal of Banking & Finance*, *27*(9), 1667–1703.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Buetow, G. W., & Henderson, B. J. (2012). An empirical analysis of exchange-traded funds. *Journal of Portfolio Management*, *38*(4), 112–127.

Burer, S., & Letchford, A. N. (2012). Non-convex mixed-integer nonlinear programming: A survey. *Surveys in Operations Research and Management Science*, *17*(2), 97–106.

Bussieck, M. R., Drud, A. S., & Meeraus, A. (2003). MINLPLib - A collection of test models for mixed-integer nonlinear programming. *Informs Journal on Computing*, *15*(1), 114–119.

Bussieck, M. R., & Vigerske, S. (2011). MINLP solver software. in J. J. Cochran, L. A. Cox Jr., P. Keskinocak, J. P. Kharoufeh and J. C. Smith (editors), Wiley Encyclopaedia of Operations Research and Management Science, Wiley, New York. Available from http://www2.mathematik.hu-berlin.de/∼stefan/minlpsoft.pdf, last accessed September 19th 2013.

Canakgoz, N., & Beasley, J. (2009). Mixed-integer programming approaches for index tracking and enhanced indexation. *European Journal of Operational Research*, *196*(1), 384–399.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*, 57–82.

Champely, S. (2013). Package 'pwr'. Available from http://cran.r-project.org/web/packages/pwr/pwr.pdf, last accessed July 30th 2013.

Charupat, N., & Miu, P. (2011). The pricing and performance of leveraged exchange-traded funds. *Journal of Banking & Finance*, *35*(4), 966–977.

Chen, C., & Kwon, R. H. (2012). Robust portfolio selection for index tracking. *Computers & Operations Research*, *39*(4), 829–837.

Cheng, M., & Madhavan, A. (2009). The dynamics of leveraged and inverse exchange-traded funds. *Journal of Investment Management*, *7*(4), 43–62.

Chiam, S. C., Tan, K. C., & Mamun, A. A. (2013). Dynamic index tracking via multi-objective evolutionary algorithm. *Applied Soft Computing*, *13*(7), 3392–3408.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ:Lawrence Erlbaum.

Cohen, L., Diether, K. B., & Malloy, C. J. (2007). Supply and demand shifts in the shorting market. *Journal of Finance*, *62*(5), 2061–2096.

Connor, G., & Lasarte, T. (2010). An introduction of hedge fund strategies. Tech. rep., London School of Economics, http://www.iam.uk.com/Portals/0/pdf/lse-publications/An-Introduction-to-Hedge-Fund-Strategies.pdf, last accessed November 4th 2013.

CPLEX Optimizer (2013). IBM. Available from http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/, last accessed August 28th 2013.

De Jong, J. C., & Rhee, S. G. (2008). Abnormal returns with momentum/contrarian strategies using exchange-traded funds. *Journal of Asset Management*, *9*(4), 289–299.

De Miguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, *22*(5), 1915–1953.

Dempster, M. A. H., Germano, M., Medova, E. A., Rietbergen, M. I., Sandrini, F., & Scrowston, M. (2006). Managing guarantees: Providing protection but maintaining access to potentially higher returns. *The Journal of Portfolio Management*, *32*(2), 51–61.

Dempster, M. A. H., Germano, M., Medova, E. A., Rietbergen, M. I., Sandrini, F., & Scrowston, M. (2007). Designing minimum guaranteed return funds. *Quantitative Finance*, *7*(2), 245–256.

Dert, C., & Oldenkamp, B. (2000). Optimal guaranteed return portfolios and the casino effect. *Operations Research*, *48*(5), 768–775.

Deville, L. (2008). Exchange traded funds: History, trading, and research. *Handbook of Financial Engineering, Springer Optimization and Its Applications, C. Zopounidis, M. Doumpos and P.M. Pardalos (editors)*, *18*, 67–97.

Egan, W. J. (2007). The distribution of S&P 500 index returns. Unpublished Manuscript, Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=955639, last accessed October 31st 2013.

Elston, F., & Choi, D. (2009). Inverse ETFs. In *Proceedings of the Academy of Accounting and Financial Studies*, vol. 14, (pp. 5–9).

Elton, E., Gruber, M., Brown, S., & Goetzmann, W. (2007). *Modern portfolio theory and investment analysis*. New York: Wiley.

Elton, E. J., Gruber, M. J., Comer, G., & Li, K. (2002). Spiders: Where are the bugs? *Journal of Business*, *75*(3), 453–472.

Engle, R., & Sarkar, D. (2006). Premium-discounts and exchange traded funds. *The Journal of Derivatives*, *13*(4), 27–45.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, *51*(1), 55–84.

Financial Stability Board (2011). Potential financial stability issues arising from recent trends in exchange-traded funds (ETFs). Basel, Switzerland. Available from http://www.financialstabilityboard.org/publications/r_110412b.pdf, last accessed October 25th 2013.

Freedman, D. (1972). *Approximating countable Markov chains*. Holden-Day.

Ganesan, G. (2011). A subspace approach to portfolio analysis: A focus on equity investments. *IEEE Signal Processing Magazine*, *28*(5), 49–60.

Garcia, F., Guijarro, F., & Moya, I. (2011). The curvature of the tracking frontier: A new criterion for the partial index tracking problem. *Mathematical and Computer Modelling*, *54*(7–8), 1781–1784.

Gastineau, G. L. (2001). Exchange-traded funds: An introduction. And further likely evolution. *Journal of Portfolio Management*, *27*(3), 88–96.

Gastineau, G. L. (2004). The benchmark index ETF performance problem: A simple solution. *Journal of Portfolio Management*, *30*(2), 96–103.

Gastineau, G. L. (2010). *The exchange-traded funds manual (2nd edition)*. Wiley Finance.

Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, *19*(3), 797–827.

Giese, G. (2010). On the risk-return profile of leveraged and inverse ETFs. *Journal of Asset Management*, *11*(4), 219–228.

Guastaroba, G., & Speranza, M. G. (2012). Kernel search: An application to the index tracking problem. *European Journal of Operational Research*, *217*(1), 54–68.

Haga, R., & Lindset, S. (2012). Understanding bull and bear ETFs. *The European Journal of Finance*, *18*(2), 149–165.

Haugh, M. B. (2011). A note on constant proportion trading strategies. *Operations Research Letters*, *39*(3), 172–179.

Herzog, F., Dondi, G., Keel, S., Schumani, L. M., & Geering, H. P. (2007). Solving ALM problems via sequential stochastic programming. *Quantitative Finance*, *7*(2), 231–244.

Hlawitschka, W., & Tucker, M. (2008). Utility comparison between security selectors, asset allocators and equally weighted portfolios within a selected ETF universe. *Journal of Asset Management*, *9*(1), 67–72.

Hutchinson, T. P., & Lai, C. D. (1990). *Continuous bivariate distributions, emphasising applications*. Rumsby Scientific Publishing.

IndexUniverse (2011). ETF briefing book: Key issues & supporting data. Available from http://www.indexuniverse.com/docs/ETFBriefingBook/IndexUniverseETFBriefing Book101811.pdf, last accessed October 25th 2013.

Investment Company Institute (2012). Chapter 3: Exchange-traded funds. Available from http://www.icifactbook.org/fb_ch3.html, last accessed October 25th 2013.

Jacobs, B. I., & Levy, K. N. (2007). 20 myths about enhanced active 120-20 strategies. *Financial Analysts Journal*, *63*(4), 19–26.

Jarrow, R. A. (2010). Understanding the risk of leveraged ETFs. *Finance Research Letters*, *7*(3), 135–139.

Jawadi, F., & Khanniche, S. (2012). Are hedge fund clones attractive financial products for investors? *Applied Economics Letters*, *19*(8), 739–743.

Joe, H. (1997). *A critique of dependence concepts*. London: Chapman & Hall.

Johnson, W. F. (2009). Tracking errors of exchange traded funds. *Journal of Asset Management*, *10*(4), 253–262.

Jolliffe, I. T. (1986). *Principal component analysis*. Springer Verlag.

Kayali, M. M., & Ozkan, N. (2007). Pricing efficiency of exchange traded funds in Turkey: Early evidence from the Dow Jones Istanbul 20. *International Research Journal of Finance and Economics*, *10*, 14–23.

Khandani, A. E., & Lo, A. W. (2007). What happened to the quants in August 2007? *Journal of Investment Management*, *5*(4), 5–54.

Khandani, A. E., & Lo, A. W. (2011). What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Financial Markets*, *14*(1), 1–46.

Korn, R. (2005). Optimal portfolios with a positive lower bound on final wealth. *Quantitative Finance*, *5*(3), 315–321.

Kostovetsky, L. (2003). Index mutual funds and exchange-traded funds: A comparison of two methods of passive investment. *Journal of Portfolio Management*, *29*(4), 80–92.

Krink, T., Mittnik, S., & Paterlini, S. (2009). Differential evolution and combinatorial search for constrained index-tracking. *Annals of Operations Research*, *172*(1), 153–176.

Kupiec, P. H. (1990). A survey of exchange-traded basket instruments. *Journal of Financial Services Research*, *4*(3), 175–190.

Kwan, C. C. Y. (1999). A note on market-neutral portfolio selection. *Journal of Banking & Finance*, *23*(5), 773–799.

Lejeune, M. A. (2011). A VaR Black–Litterman model for the construction of absolute return fund-of-funds. *Quantitative Finance*, *11*(10), 1489–1501.

Leyffer, S., Linderoth, J., Luedtke, J., Mahajan, A., & Munson, T. (2013). Minotaur solver. Available from http://wiki.mcs.anl.gov/minotaur/, last accessed September 19th 2013.

Lin, V., & Mackintosh, P. (2010). ETF mythbuster: Tracking down the truth. *The Journal of Index Investing*, *1*(1), 95–106.

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

Little, P. K. (2010). Inverse and leveraged ETFs: Not your father's ETF. *The Journal of Index Investing*, *1*(1), 83–89.

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, *3*(2), 175–205.

Lo, A. W., & Patel, P. N. (2008). 130/30: The new long–only. *Journal of Portfolio Management*, *34*(2), 12–38.

Ma, Y., MacLean, L., Xu, K., & Zhao, Y. (2011). A portfolio optimization model with regime-switching risk factors for sector exchange traded funds. *Pacific Journal of Optimization*, *7*(2), 281–296.

Madhavan, A. (2012). Exchange-traded funds, market structure, and the flash crash. *Financial Analysts Journal*, *68*(4), 20–35.

Maister, D., Schorr, M., & Perlman, D. (2009). ETFs: Average tracking error was well contained in 2008. *ETFs and Indexing*, *1*, 108–125.

Maister, D., Schorr, M., Perlman, D., & Minar, S. (2010). Exchange-traded funds: Average tracking error rose meaningfully in 2009. *The Journal of Index Investing*, *1*(1), 132–163.

Mariani, M. C., Libbin, J. D., Martin, K. J., Ncheuguim, E., Varela, M. P. B., Mani, V. K., Erickson, C. A., & Valles-Rosales, D. J. (2009). Levy models and long correlations applied to the study of exchange traded funds. *International Journal of Computer Mathematics*, *86*(6), 1040–1053.

Maringer, D. (2008). Constrained index tracking under loss aversion using differential evolution. *Studies in Computational Intelligence*, *100*, 7–24.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.

Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments*. New York: Wiley.

Meade, N., & Beasley, J. E. (2011). Detection of momentum effects using an index outperformance strategy. *Quantitative Finance*, *11*(2), 313–326.

Mezali, H., & Beasley, J. E. (2013). Quantile regression for index tracking and enhanced indexation. *Journal of the Operational Research Society*, *64*, 1676–1692.

MINLP Library (2013). Available at http://www.gamsworld.org/minlp/minlplib.htm, last accessed September 19th 2013.

Mittelmann, H. D. (2012). Performance of commercial and noncommercial optimization software. Presented at INFORMS 2012, Phoenix, Arizona, USA. Available from http://plato.asu.edu/talks/phoenix.pdf, last accessed September 19th 2013.

Nishiyama, N. (2001). One idea of portfolio risk control for absolute return strategy risk adjustments by signals from correlation behavior. *Physica A*, *301*(1–4), 457–472.

Pai, G. A. V., & Michel, T. (2012). Differential evolution based optimization of risk budgeted equity market neutral portfolios. In *In Proceedings of WCCI 2012, IEEE World Congress on Computational Intelligence (IEEE–WCCI)*. 345 East 47th Street, New York, NY 10017 USA: IEEE.

Patton, A. J. (2009). Are "market neutral" hedge funds really market neutral? *Review of Financial Studies*, *22*(7), 2495–2530.

Pole, A. (2008). *Statistical arbitrage: Algorithmic trading insights and techniques*. Wiley Finance.

Poterba, J. M., & Shoven, J. B. (2002). Exchange-traded funds: A new investment option for taxable investors. *American Economic Review*, *92*(2), 422–427.

Ramaswamy, S. (2011). Market structures and systemic risks of exchange-traded funds, Working paper No 343, Bank for International Settlements, Basel, Switzerland. Available from http://www.bis.org/publ/work343.pdf, last accessed October 25th 2013.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional Value-at-Risk. *Journal of Risk*, *2*, 21–41.

Rollenhagen, K. J. (2009). Pitfalls and potential: An attempt to demystify leveraged & inverse exchange-traded index funds. *ETFs and Indexing*, *1*, 105–107.

Rompotis, G. G. (2008). Performance and trading characteristics of German passively managed ETFs. *International Research Journal of Finance and Economics*, *15*, 218–231.

Rompotis, G. G. (2009). Interfamily competition on index tracking: The case of the vanguard ETFs and index funds. *Journal of Asset Management*, *10*(4), 263–278.

Rompotis, G. G. (2011a). A study on the third-generation exchange-traded funds: The case of short ETFs. *The Journal of Index Investing*, *2*(1), 25–43.

Rompotis, G. G. (2011b). Predictable patterns in ETFs' return and tracking error. *Studies in Economics and Finance*, *28*(1), 14–35.

Rompotis, G. G. (2011c). The performance of actively managed exchange-traded funds. *The Journal of Index Investing*, *1*(4), 53–65.

Rompotis, G. G. (2012). A survey on leveraged and inverse exchange-traded funds. *The Journal of Index Investing*, *2*(4), 84–95.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341–360.

Ruiz-Torrubiano, R., & Suarez, A. (2009). A hybrid optimization approach to index tracking. *Annals of Operations Research*, *166*(1), 57–71.

Sabbaghi, O. (2011). The behavior of green exchange-traded funds. *Managerial Finance*, *37*(5), 426–441.

Schmidhammer, C., Lobe, S., & Roder, K. (2011). Intraday pricing of ETFs and certificates replicating the German DAX index. *Review of Managerial Science*, *5*(4), 337–351.

Scozzari, A., Tardella, F., Paterlini, S., & Krink, T. (2013). Exact and heuristic approaches for the index tracking problem with UCITS constraints. *Annals of Operations Research*, *205*(1), 235–250.

Sharifzadeh, M., & Hojat, S. (2012). An analytical performance comparison of exchange-traded funds with index funds: 2002-2010. *Journal of Asset Management*, *13*(2), 196–209.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), pp. 425–442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, *39*(1), 119–138.

Sharpe, W. F. (1975). Adjusting for risk in portfolio performance measurement. *Journal of Portfolio Management*, *1*(2), Winter:29–34.

Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, *21*(1), Fall:49–58.

Shin, S., & Soydemir, G. (2010). Exchange-traded funds, persistence in tracking errors and information dissemination. *Journal of Multinational Financial Management*, *20*(4–5), 214–234.

Sortino, F., & van der Meer, R. (1991). Downside risk. *The Journal of Portfolio Management*, *17*(4), 27–31.

Storn, R., & Price, K. (1997). Differential evolution - a simple and efficient heuristic for global optimisation over continuous spaces. *Journal of Global Optimization*, *11*(4), 341–359.

Székely, G. J., & Rizzo, M. L. (2009). Brownian distance covariance. *The Annals of Applied Statistics*, *3*(4), 1236–1265.

Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, *35*(6), 2769–2794.

Thomson Reuters Datastream (2011). Private communication.

Thomson Reuters Datastream (2013). Available from http://online.thomsonreuters.com/Datastream/, last accessed August 28th 2013.

Till, H., & Eagleeye, J. (2003). Traditional investment versus absolute return programmes. *Quantitative Finance*, *3*(3), 42–48.

Tol, R., & Wanningen, C. (2009). On the performance of extended alpha (130/30) versus long-only. *Journal of Portfolio Management*, *35*(3), 51–60.

Tol, R., & Wanningen, C. (2011). 130/30: By how much will the information ratio improve? *Journal of Portfolio Management*, *37*(3), 62–69.

Treynor, J. (1961). Towards a theory of market value of risky assets. Unpublished manuscript.

US Senate (2010). Findings regarding the market events of May 6, 2010: Report of the staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. Available from http://banking.senate.gov/public/_files/StaffFindingsregardingtheEventsofMay6.pdf, last accessed October 25th 2013.

van Montfort, K., Visser, E., & van Draat, L. F. (2008). Index tracking by means of optimized sampling. *Journal of Portfolio Management*, *34*(2), 143–151.

Wang, M., Xu, C., Xu, F., & Xue, H. (2012). A mixed 0-1 LP for index tracking problem with CVar risk constraints. *Annals of Operations Research*, *196*(1), 591–609.

Waring, M. B., & Siegel, L. B. (2006). The myth of the absolute-return investor. *Financial Analysts Journal*, *62*(2), 14–21.

Wilmott, P. (1998). *Derivatives: The theory and practice of financial engineering*. New York: Wiley.

Wong, K. H. Y., & Shum, W. C. (2010). Exchange-traded funds in bullish and bearish markets. *Applied Economics Letters*, *17*(16), 1615–1624.

Zymler, S., Rustem, B., & Kuhn, D. (2011). Robust portfolio optimization with derivative insurance guarantees. *European Journal of Operational Research*, *210*(2), 410–424.