THE OPTIMAL ASSET ALLOCATION FOR SOUTH AFRICAN REAL RETURN INVESTORS

Barry van Zyl

Dissertation presented for the degree of
Master of Statistical Science
in the Department of Statistical Sciences at the University of Cape Town, Cape Town,
November 2018

Supervisor: Prof. David Bradfield

To Stefanie,

thank you for your love and support.

ABSTRACT

This research aims to establish the optimal asset allocations for targeting specific real returns over short, medium and long-term investment horizons. The joint returns are modelled with data-centric methods that are empirical and non-parametric in nature, and are able to capture the dependencies of returns over time.

The asset classes that are considered are South African (SA) equities, SA bonds, SA cash, SA property, global equities, global bonds, global cash, and global property. The returns of each asset class are modelled, each class with its own empirical distribution based on monthly returns from 1972 to 2017.

The monthly returns are grouped in a block of rolling periods of varying block lengths in order to attempt to capture dependencies across time. These blocks of data are resampled in order to simulate the distributions of returns of portfolios with their own unique empirical distribution.

The optimal portfolios are derived using a genetic algorithm, showcasing how these extremely versatile optimisation tools can be used in combination with resampling methods to find the optimal portfolio for virtually any criterion. A comparison is also made to the traditional mean-variance optimal portfolios, yielding an estimate of the bias in mean-variance optimisation's (MVO) optimal weights.

It is investigated how these optimal portfolios are influenced by the choice of risk criterion and investment horizon. The effect of the most important and consequential nuisance parameter in this research's model, the block length, is discussed. The relationships established between the characteristics of optimal portfolios and investment horizon and risk criterion and the comparisons with classic MVO should be of interest to investors and investment professionals alike.

Economic and market regimes are "identified" on the basis of economic and market data, consequently the resampling probabilities will be unequal. The optimal weights conditional on regimes are derived. Both static and changing regimes are considered.

Lastly, an out-of-sample backtest of the performance of the optimal portfolios conditional on the regime across time at six month intervals is conducted from 1983 to 2017. It shows that out of the three block lengths tested for a single investment horizon of 36 months, a block length of 24 months yielded the best overall risk-adjusted performance, on average. Conditioning for regimes is shown to generally outperform the unconditional approach. The improvements are marginal and further research is recommended to investigate the performance for longer investment horizons and other values of the two tuning parameters, block length and tactical pressure.

The higher level aim of this work is to present a broad sense of how data-driven nonparametric methods can be used in conjunction with metaheuristic procedures. The objective of combining these techniques is to find optimal portfolios under very general conditions and with very few assumptions regarding the underlying distributions.

TABLE OF CONTENTS

1. INTRODUCTION

Mean-variance optimisation

In 1952 Harry Markowitz (Markowitz, 1952) presented the first mathematical framework that explicitly allowed for the trade-off between returns and risk in the investment portfolio decision. The key insight of his theory is that the assets should not be considered in isolation, but instead as the building blocks of an overall portfolio that should simultaneously allocate to all available assets. Even though Markowitz's insight is still considered a breakthrough, the logic now seems deceptively simple: An investor should choose a portfolio of assets with the minimum possible combined risk that satisfies a certain return threshold; or alternatively, the portfolio of assets with the highest combined return given a certain maximum risk threshold.

The popularity of the method is arguably driven to a large extent by its convenience. It enables the rational combination of varied assets, each potentially of a very different nature – all that is required is a historical return series for every asset in the opportunity set.

However, the method also has its shortcomings, as acknowledged very early on by Markowitz himself. Mean-variance optimisation (MVO) assumes that returns are independently and identically distributed over time. As a result MVO is a one-period only optimisation that does not allow for the incorporation of changing processes governing asset returns, or the investment horizon of the investor.

Further, since only the mean and variance of the optimal portfolio is considered, it implicitly assumes that either assets are normally distributed, or investors are indifferent to the implications of potentially non-normally distributed returns.

These shortcomings pose several questions to the rational investor: Are returns normally distributed, and what if they are not? If returns are, for example, not symmetric, is it appropriate that MVO assumes that investors are equally averse to returns above the mean than returns below the mean? Are the returns of asset classes the result of stationary processes without interdependencies across time? How can investment horizons that encompass several non-identically distributed periods be incorporated into the investment process to arrive at multi-period optimal portfolios?

The non-normality and non-stationarity of the joint returns of asset classes lead this study to consider alternative distributions to the normal distributions and alternative measures of risk, as well as the possibility of a time-varying opportunity set.

Strategic and tactical asset allocation

Asset allocation, as it is used in this research and in the portfolio management industry, refers to the allocation of assets to a relatively small number of "asset classes".

Balanced funds in South Africa, for example, typically comprise South African (SA) equities, SA bonds, SA cash, SA property, and global equities, global bonds, global cash and global property. At any one point in time some or all of these buckets of assets may be present in any given balanced fund. Each of these buckets consists of a diversified portfolio of securities, which could be either a passive investment (e.g. index tracking) or an active investment. In either case, each of these portfolios is typically managed against a market benchmark, and exhibits some degree of similarity to the market benchmark in terms of its performance. Table 1 shows typical benchmarks used in South Africa to proxy the performance of the various asset classes:

Asset class	**Market benchmark or proxy**
SA equities	JSE Shareholder Weighted Index (SWIX)
SA bonds	JSE All Bond Index (ALBI)
SA cash	STEFI (3M)
SA property	JSE SA Listed Property
Global equities	MSCI AC World
Global bonds	Barclays Global Aggregate
Global cash	US, UK, EURO LIBID rates
Global property	UBS Global Property Index

Table 1: Typical market proxies

If the balanced fund is an active fund, security selection (i.e. the securities selected to be held and the proportion allocated to each) will take place within each portfolio in an attempt to outperform each respective benchmark. However, the asset allocation – i.e. how much is allocated to each underlying specialist asset class portfolio – is a separate and distinct decision from the security selection.

Asset allocation is often separated into a long-term or strategic asset allocation (SAA) decision and a shorter term or tactical asset allocation (TAA) decision. The strategic asset allocation is, in essence, the long-term asset allocation that, in the absence of tactical asset allocation, is deemed most likely to achieve the objectives of the fund. It is also the default allocation if the asset manager has no tactical views, or the reference asset allocation relative to which the tactical asset allocation is positioned. In other words, the TAA often tilts away from an explicit SAA to account for the shorter term views of the portfolio manager.

Very often the objective of a balanced fund, and hence the SAA, is to achieve a real return (return in excess of inflation) over a certain period of time (often referred to as the investment horizon). Many of the investors in real return targeting balanced funds are advised by financial planners to invest in these funds on the basis of the specific circumstances of the investor. The financial advisor often performs a detailed expected future cash flow (or asset-liability) analysis to determine the client's required real return to make ends meet.

The SAA and TAA decisions are tremendously decisive in the long term as to whether the real return objective of a fund will be met, and hence whether the financial plan of an investor will come to bear. The consequences of a financial plan not going according to plan due to underperformance of the required real return can potentially be catastrophic from the point of view of the individual involved. In an oft-cited paper "Does asset allocation explain 40, 90 or 100 Percent of Performance", the authors, Ibbotson & Kaplan (2000), come to the conclusion that the policy (or target) asset allocation over time explains 40% of the variation of returns among funds[1], 90% of the variability of a fund's returns over time[2] and more than a 100% of the level of returns.[3] Having an appropriate asset allocation across all life-stages is therefore clearly of great importance to individual investors.

Ideally SAAs should remain unchanged over long periods of time. Though there is little research on the practical considerations around changing a fund's SAA in a real-life setting, personal experience leads us to posit that such attempts may be met with resistance from stakeholders. Consider a hypothetical scenario of a fund with an SAA. As the SAA will, as shown by Ibbotson and Kaplan, have an important effect on future performance, it is very possible that it is an essential input in determining the risk profile and expected returns of a fund, which in turn could appear on marketing material (or perhaps even the mandate or trust deed of a fund), and an important point of discussion in correspondence between the asset managers, financial advisors and investors. The SAA in and of itself may be an important consideration when selecting a fund. It therefore stands to reason that changing the SAA could have material repercussions for the risk and return profile of a fund and may invalidate or at least call into question earlier discussions between stakeholders, especially within the financial planning exercise. It is not difficult to imagine a situation in which changing the SAA of a fund means that it will no longer be appropriate (or at least perceived to be appropriate) for the underlying investor.

It therefore also stands to reason that changing an SAA may be met with resistance, and may therefore remain unchanged for long periods of time. This, in turn, implies that it could have repercussions on long-term performance (given that managers tend to be acutely aware of performance benchmarks and objectives, and thus tend to be very cognisant of their tilts relative to SAAs). Although the literature has very little to say about the management of SAAs in practice, at the very least it seems clear that, all else being equal, the less frequently an allocation that is treated and sold as a long-term benchmark around which a fund will position itself changes, the better. Further, the less frequently an SAA changes, the more pronounced its impact will be on performance.

The optimal asset allocations derived in this work will speak to the questions of optimal strategic (long-term, independent of current market and economic conditions) as well as tactical asset allocations (short to medium-term, conditional on current conditions, regime-based).

---

[1] The percentage difference between two funds' performance explained by different target allocations.
[2] The percentage variability over time of a single fund's returns that is explained by the target allocation.
[3] The ratio of the target asset allocation over time's return relative to the actual fund return.

1.1. Thesis statement and outline

This thesis will outline and implement an empirical approach to determine the optimal strategic and tactical allocations for a South African investor with a real return objective and an explicit investment horizon. Non-normality and non-stationarity of returns will be allowed for in several ways:

(1) via the empirical distributions of asset allocations;
(2) via conditioning on regimes (effectively skewing the resampling probabilities in the empirical distribution);
(3) via the resampling *blocks* of returns instead of single months (thereby preserving or partially preserving time-dependency structures in the data and incorporating the consideration of the periodicity or frequency of return units in the analysis); and
(4) via the use of various risk criteria (each with its own implied utility function), that specifically take into account the investment horizon in their definitions (as opposed to simply the standard deviation of monthly returns, as is typically the case).

The distribution of each and every portfolio of assets will be modelled by its own empirical distribution, i.e. the shape of the distribution in each case is not arbitrarily imposed but rather dictated by the data itself.

The methodological choices in points (3) and (4) above make it possible to investigate the role of the length of the investment horizon in the optimal portfolio. The effects on optimal weights of both time-dependencies of asset returns as well as the use of different risk criteria are quantified.

The periodicity of the data will be dependent on the length of the investment horizon, in order to explicitly model the effect of time-dependencies in returns (e.g. the short-term momentum and long-term mean reversion exhibited by equity returns), and to contrast the results to the predominant methodology in practice and the literature (i.e. MVO based on an assumption that monthly returns are identically and independently distributed).

The optimal portfolios belonging to a large number of problem settings, each with its own unique set of parameters (e.g. required return, investment horizon and risk criterion), will be found with the help of the bespoke implementation of a metaheuristic technique called the genetic algorithm. The effect of each parameter on the optimal portfolio will be discussed in some detail, and the results will be contrasted with more traditional MVO approaches.

The outline of this dissertation will be as follows:

In **Chapter 2** literature on the distribution of asset returns and optimal portfolios incorporating regimes is reviewed. Although this study's approach to modelling asset returns is rather novel, the strengths and weaknesses of existing approaches will serve as an important counterpoint to the methodology employed in this study.

In **Chapter 3** the distributional assumptions typically made in the literature and in practice are discussed. Other key terms and concepts employed to define the problem settings within which optimal portfolios will be derived are then introduced: the required return, investment horizon, monthly standard deviation, standard deviation over horizon, probability of success, average shortfall, downside deviation and maximum drawdown, among other terms.

Chapters 4 (DATA AND METHODOLOGY) and 5 (RESULTS) are the most substantive and important chapters of this work.

Chapter 4 starts by introducing the data that will be used to model historical returns in Chapter 5. There are two major categories of data: (1) monthly asset class returns; and (2) regime-identifying variables (the latter consisting of both market and economic variables). Next the return, risk and time-dependencies of asset returns are reported. The time-dependencies of returns will be referred to very frequently in the results discussed in Chapter 5. Section 4.2 defines various concepts that will form the components of this study's returns distribution models: the empirical distribution, the block bootstrap and the regime-classification methodology used in this study.

Section 4.3 formally specifies the full models. The components of Section 4.2 are combined to define the final models: one regime-ignorant model and three regime-cognisant models.

The technical aspects of the optimisation procedure that searches the solution space for optimal portfolios are, for the sake of brevity and flow, relegated to Appendix 7.1. This section introduces the reader to genetic algorithms and then describes this study's bespoke genetic algorithm. While this appendix item can be skipped without sacrificing understanding of later sections, it is quite detailed as it is an important component of this research as our novel returns distribution model requires a novel, technically challenging approach to finding optimal portfolios.

Chapter 5 discusses the optimal portfolios found by the genetic algorithm. The effects of the risk criterion, investment horizon, and expected block length on the optimal allocation in the case of regime-ignorant optimal portfolios are reported in Section 5.1. There is also a comparison of the optimal portfolios found to those of traditional MVO based on monthly data.

In the regime-cognisant results (Section 5.2), the focus is initially narrowed to a single block length and investment horizon in Sections 5.2.1 and 5.2.2, for the sake of brevity. Extensive out-of-sample backtesting is performed for the chosen setting. In Section 5.2.3, we widen this regime-cognisant setting to include several risk criteria and block lengths.

Chapter 6 summarises the conclusions and recommendations for further research.

2. LITERATURE REVIEW

Non-normality of returns

Markowitz's theory, in its most naïve and common form, assumes that returns are jointly normally distributed. One implication of this assumption is that all periods are identically and independently distributed (i.i.d.) irrespective of the periodicity. It has long been understood that this assumption is not strictly realistic. It is trivial to note, for example, that the returns on cash instruments generally display high levels of autocorrelation for periods closer in proximity (Ang & Bekaert, 2004:1). A very commonly cited example in the literature is the persistent change in the distribution of returns of assets during and following the 2008 financial crisis (Bae, Kim & Mulvey, 2014:453; and Ang & Timmerman, 2012:1) and other extreme market scenarios.

Asset returns are also well-known to exhibit certain *stylised facts*, a collective term given to distributional properties (typically incompatible with the joint normal distribution) that can be observed in returns. For example, Cont (2001) examined a wide range of assets and reported, among other traits, fat tails, asymmetry, volatility clustering, some evidence for autocorrelation for weekly to monthly returns, and clear autocorrelation on nonlinear functions of the return (e.g. on squared or absolute returns). These stylised facts can themselves to an extent also be time-varying, as noted by for example Aït-Sahalia & Brandt (2001:1333) in the case of higher order moments.

Stock returns are often characterised as exhibiting momentum over shorter periods and mean reversion over longer periods. Several seminal papers in the late 1980s queried the efficient market hypothesis by investigating these two purported phenomena in stock returns. Fama & French (1988) asserted evidence that mean reversion accounted for in the region of 40% of the variation in three to five year returns of United States (US) stocks. Poterba & Summers (1986) examined stock returns from as early as 1871 to 1986 from 17 different countries and reported shorter term (less than 12 months) momentum and longer term (longer than at least 24 months) mean reversion that explain as much as half of the variance in monthly returns.

Carhart's (1997) suggestion that Fama & French's (1993) popular three factor pricing model for equities size, value, and growth be extended to include a fourth factor, namely one-year momentum, has been widely adopted by theorists and practitioners. More recently studies have reported mixed findings on whether long-run mean reversion exists, at least in part due to the small number of independent sample points that are available for periods of time that are several years in length (Spierdijk, Bikker & Van den Hoek, 2012:5). Evidence for momentum is, however, stronger, as reported by Fama & French (2012) for one-year momentum in North America, Europe and Asia-Pacific stocks.

Optimal portfolios under regimes

Overview

Most generally, the theory or model underlying the optimal portfolio should take account of a time-varying opportunity set. However, *cyclical* components to businesses and economies have also long been thought to exist in some sense, and it is natural to suspect that these ebbs and flows have direct or indirect repercussions on markets and asset returns. One of the earliest detailed analytical treatments on this topic is an exposition on periodic economic crises by Jean Charles Léonard Simonde de Sismondi (1819). If such cycles occur, there may exist exploitable patterns, and the need for dynamic asset allocation becomes apparent: depending on where in a cycle the economy finds itself, the optimal portfolio will naturally change – if the current position in the cycle and the returns distributions conditional on the current position in the cycle can be distinguished.

There is a large body of work on the modelling of the distributions of cyclical economic variables, and the returns distributions of asset classes in the presence of such cycles. Typically it is assumed that there is a finite number of regimes, usually between two and six. Often the regimes are noted to be associated with "good" states, "bad" states, and states intermediate to these two states. Regimes are usually thought of as discrete categories that correspond to economic or financial fundamental phenomena (e.g. growth, recovery, contraction, etc.), but sometimes the states are continuous (e.g. Fu, Wei & Yang, 2014). Regimes may alternatively also correspond to periods in-between various secular changes, for example regulatory changes, or changes in economic, fiscal, or monetary policy (Ang & Timmerman, 2012:1).

The number of states chosen is often arbitrary or not strictly dictated by the data, or the number of states is sometimes limited purely by the number of data points available, especially the number of data points that can be assigned to each state. The possible loss of realism due to the enforcing of a small number of states by the limited available data is usually not interrogated in any detail, or even briefly discussed. In some cases the Bayesian Information Criterion (or a similar criterion) is used, along with heuristic arguments, to determine the number of states (e.g. Bae, Kim & Mulvey, 2014). Some authors (e.g. Ang & Timmerman, 2012:4) argue that the number of regimes should as far as possible be based on economic arguments.

The defined regimes are usually assumed to be essentially the same over time (i.e. cyclical), only changing marginally as the dataset expands over time. Alternatively, a regime's characteristics may change more abruptly due to a weighing scheme on past data, for example an exponentially declining weight scheme based on the recentness of data (e.g. Nystrup, Madsen & Lindström, 2017).

For the sake of simplicity, regime-cognisant models sometimes assume that the current regime is observable with certainty and will remain static. In this scenario the optimal allocation usually depends only on this observed, prevailing regime.

Another possibility is to not assume the current regime is known with certainty, but rather to assign a probability of it being any one of the possible regimes. If no future regime-changes are allowed for, the optimal portfolio is essentially a weighted average of the optimal portfolio in each regime, where the weight to each is determined by the aforementioned probabilities.

Most generally, the current regime is not known with certainty, and the transition between regimes is explicitly modelled, typically with a Markov chain, allowing for an optimal portfolio that may take into account several future regime changes. In this more realistic scenario, optimal allocations take into account uncertainty about prevailing regimes and future regimes.

Whether future regime changes are modelled or not, the optimal asset allocation should change over time when the state changes (or the probabilities of being in the various regimes change) as new data comes to light. In reality the transitions between different regimes may or may not be predictable, but the very existence of regimes can parsimoniously explain a large number of phenomena, such as stylised facts of asset distributions (e.g. skewness, kurtosis and time-varying correlations), and non-monotonic risk/return relationships (Ang & Timmerman, 2012:1).

Conditional on the regime, the distribution of returns is often assumed to be normally distributed, but with different joint means, variances and correlations in each regime. If the probability of transitioning between states are constant across time, the unconditional distribution is a mixture of normal distributions weighted by these probabilities (Ang & Bekaert, 2002; Guidolin & Timmerman, 2005; and Timmerman, 2000). This scheme can reproduce fat tails and persistent volatility (or volatility clustering), amongst other stylised effects.

Return distribution regimes may be identified *endogenously* to returns data, i.e. via the asset returns themselves (e.g. Chow, Jacquier, Kritzman & Lowry, 1999; Ang & Bekaert, 2004; Honda, 2003; Guidolin & Timmerman, 2005; and Nystrup, 2014), or *exogenously* to the returns data – i.e. on macroeconomic or other non-returns data (e.g. Turner & Han, 2009; and Munro & Silberman, 2008) – or some combination of endogenous and exogenous variables (e.g. Kritzman, Page & Turkington, 2012).

When time periods are exogenously classified as belonging to a number of groupings or clusters, this classification is usually assumed to be known with certainty. The classification may be performed on an ad hoc visual studying of economic data (e.g. Munro & Silberman, 2008), or employ more sophisticated multivariate techniques such as principal components (e.g. Kondlo, 2016) or Mahalanobis distance (e.g. Chow, Jacquier, Kritzman, & Lowry, 1999).

The loss of realism due to the assumption of deterministic states is not typically considered in detail. Further, when data is classified in this manner, the implicit assumption is made that the returns of any one period are completely irrelevant to the distribution of returns in a different period. However, if the current regime is still known only in probability, the overall optimal portfolio may still essentially be a weighted average of the optimal portfolios conditional on the regimes – i.e. a dependent on all the returns data. Alternatively, the optimal portfolio may refer to expected returns and a covariance matrix that is a weighted (with the probability of being in the relevant state) average of the expected returns and covariances of each regime (e.g. Ang & Bekaert, 2004; and Chow, Jacquier, Kritzman, & Lowry, 1999).

In contrast, if the state at the time of analysis is assumed known and assumed to remain unchanged over the investment horizon, the optimal portfolio will not be a weighted average but will instead simply be one of the conditionally optimal portfolios. Thus, if the state is assumed known with certainty, the optimal portfolio at a given point in time will deviate more from the optimal portfolio belonging to an i.i.d. setting (Ang & Bekaert, 2002:1142)[4].

The data is in the vast majority of cases of monthly periodicity, and state changes can occur after every month-end. The shape of the distribution of returns over longer periodicities is captured (or meant to be captured) by the combination of the state changing mechanism and the conditional distributions. Nystrup (2014) is an example where daily data is used and daily regime-changes are allowed for.

Hidden Markov models

One of the more sophisticated and general approaches is to model state changes with a so-called Hidden Markov model (e.g. Ang and Bekaert, 2002; Ang & Timmerman, 2012; and Hamilton, 1989). The great strength of this model is that it *simultaneously* fits regime shifts as well as the parameters of the conditional distributions.

The state at any point in time is not observable with certainty, but rather inferred from the returns data, hence the qualifier "hidden". In some cases, once the state is inferred, it is assumed known with certainty for the sake of simplicity and for the purposes of fitting a model. More generally, the Markov process is assessed to be in any one state with some probability, at any point in time.

This model is sometimes explained by way of comparison to inferring whether a person is awake or asleep when only his heartbeat across time is known – an awake state will generally have a higher heartbeat, and an asleep state a lower heartbeat. However, one cannot know for certain the state of the person, as a nightmare during the asleep state may also cause a raised pulse, and a deeply relaxed awake person may have a very slow heartbeat, and so on (Kritzman, Page & Turkington, 2012:23).

The Hidden Markov model is typically of the first order – i.e. the probability of changing to any state is only dependent on the current state.

[4] To see this, consider that the single optimal portfolio in an i.i.d. setting is roughly speaking also a weighted average of the conditionally optimal portfolios of a regime setting (with the weights determined by the number of data points in each regime).

Fitting the model

The unobservability of regimes poses a significant technical challenge as far as the fitting of the model is concerned. The main approaches are maximum likelihood estimators in conjunction with filters (e.g. Hamilton, 1989; and Ang & Bekaert, 2004) or the expectation maximisation algorithm (e.g. Fraser, 2008), Bayesian methods such as the Gibbs sampling algorithm and Markov Chain Monte Carlo methods (e.g. Lam, So & Li, 1998) and other Monte-Carlo simulation based methods (e.g. Guidolin & Timmerman, 2005; and Honda, 2003). As already mentioned, the parameters of the conditional distributions and the probability of being in any one regime at any point in time are both simultaneously estimated and are outputs of the same procedure.

Transition probabilities

Prevailing states are generally found to be highly persistent, which increases the ability of regime-cognisant asset allocation to exploit time-varying returns – the more persistent regimes are, the more predictable the distributions of asset returns. Good or neutral states are generally more persistent than bad states. Gray (1996:41) finds in the context of short-term interest rates that both good and bad states persist with probabilities exceeding 0.9 from one month to the next. Ang & Bekaert (2002:1152) find that both good and bad states in the covariances of equities persist with monthly probabilities of more than 0.8. Ang & Bekaert (2004:16), in the context of US equities, US bonds and US cash, again find persistence probabilities in excess of 0.8. Guidolin & Timmerman (2005:7) find in the context of US equities and US cash that good and bad states persist with monthly probability of 0.95 and 0.81 respectively. Sa-Aadu, Shilling & Tiwari (2005:35) examine optimal asset allocation to US stocks, offshore stocks, US bonds and US real estate, and report that the good state persists with probability 0.79 from one month to the next, while bad state persists with probability of only 0.61. Bae, Kim & Mulvey (2014:452) estimate that in their two-state model in the context of US equities, commodities and US bonds, the good state persists (0.91) with higher probability than the bad state (0.71).

The transition probabilities themselves may also be a function of time. For example, Ang and Bekaert (2004) allow transition probabilities to depend on the interest rate, and report significant evidence that such a relationship does indeed exist.

Out-of-sample performance

Several studies report on the out-of-sample performance of their models. The performance in out-of-sample tests are generally (not surprisingly) inferior to in-sample performance (Nystrup, Hansen, Madsen & Lindström, 2015:104). Nystrup et al. reported in their study (which focussed on global equities and global bonds) that, out-of-sample, the breakeven transaction costs versus a passive strategy was 239 basis points (bps). In the presence of a dynamic asset allocation, it may be important to incorporate transaction costs in order to fairly compare an active strategy with high turnover to a passive strategy. Kritzman, Page & Turkington (2012:29) included global stocks, hedge funds, global bonds, gold, US cash and inflation-linked bonds in their study and found in out-of-sample testing that the breakeven transaction cost was 133 bps. Ang & Bekaert (2004:18) reported that out-of-sample tests delivered an outperformance relative to a non-regime dependent method of between 1.5% and 7.5% per annum, depending on the risk-aversion parameter. Guidolin & Timmerman (2007:20) performed out-of-sample forecasting experiments that, according to the authors, confirmed the importance of accounting for regimes in US stock and US bond returns.

Selected studies

This section gives a brief synopsis of the some of the most cited and influential studies involving modelling regimes. The aim is to give the reader a brief sense of how these studies vary in methodology and context, and to serve as counterpoint to the methodological choices of this research.

As early as 1969, **Merton (1969)** considered the optimal portfolio in a theoretical setting where the return generating process changes continuously over time. He noted that the returns generated under these circumstances are not i.i.d., and that the optimal portfolio could be separated into a myopic component and a hedging component. The myopic portfolio is the optimal portfolio for the one-period case, while the hedging component was so called as it hedged against unfavourable changes in the returns generating process. However, Merton's model is analytically cumbersome and the applicability to real-life situations was limited at the time (Van Wyk de Vries, Gupta & Van Eyden, 2014:2).

The majority of contemporary attempts to model financial and economic regimes can be traced back to **Hamilton's (1989)** treatment of US Gross National Product (GNP) growth. Although Hamilton's study was not in the context of asset selection or the modelling of return distributions, the methods used to model unobserved regimes proved seminal for regime-cognisant returns modelling.

Hamilton proposed a first-order Markov-switching regression with two discrete, unobserved states ("fast growth" and "slow growth"). The success of Hamilton's work owes much to his further contributions to the fitting of the parameters. In order to identify a change in regime, Hamilton devised a Bayesian "filter" (building on work by Cosslett & Lee, 1985) that is appropriate in a setting where the state is only known in probability.

Hamilton found that, rather than characterising GNP growth as a somewhat stable long-term trend with periods of slower and faster growth, his model divided the period in consideration into periods of positive and negative growth states that resembled the National Bureau of Economic Research's (NBER) dating of business cycles remarkably closely.

Chow, Jacquier, Kritzman & Lowry (1999) attempted to address the instability of the covariance matrix and the consequent effects on the appropriateness of the standard optimal portfolio at different points in time. They endogenously defined regimes on monthly asset class returns by categorising each set of monthly returns as either an outlier or not. Outlying observations were classified as belonging to a turbulent regime, while the rest belonged to a quiet regime. They simply grouped the 25% of monthly returns that were furthest from the mean according to the Mahalanobis distance together (the turbulent times regime) and the rest of the observations were grouped together (the quiet times regime). No theoretical or practical argument is given for the arbitrary choice of 25%.

They then calculated the sample covariance matrix belonging to each of these regimes. Finally, they suggested that the user of this method blends the two covariance matrices depending on his perceived probability of being in either regime. The final covariance matrix is a simple weighted average of the two regime specific covariance matrices, and is subsequently used as the input into a mean-variance optimisation. Bauer, Haerden & Molenaar (2004) implemented the ideas of Chow et al. on monthly data from 1976 to 2002 for six asset classes from the point of view of a US investor, and concluded that an investor with perfect foresight on the next month's regime would marginally outperform the standard MVO investor after taking into account transaction costs.

The work done by **Ang & Bekaert (2002)** is one of the first studies to incorporate regimes that explicitly follow a Markov chain in an asset allocation setting, albeit within the context of a very specific problem. The authors wished to address the puzzle posed by the phenomenon that investors tend to invest proportionately more in their local stock markets than in international stock markets than predicted by the prevailing standard models that assume constant volatilities and correlations. One theory often offered at the time was that the correlations between international equity returns increased during times of distress, precisely when investors needed protection, thereby diminishing the diversification benefits predicted by standard models. The authors attempted to resolve this "home bias puzzle" by investigating whether this explanation holds water.

They assumed a regime-switching data generating process with known and *observable*[5] regimes with some degree of persistence. Conditional on the current regime, the joint distribution was assumed to be normal. The conditional joint distribution one period into the future is thus a mixture of normal distributions, which is able to reproduce fat tails and persistent volatility, among other effects.

The optimal portfolio is a function of the current state (regime) as well as future states, as it must take into account the probability of switching to the next state (which depends on the current state). As a result the optimal portfolio is also a function of the investment horizon: for example, the longer the investment horizon, the higher the probability of eventually being in all states. The optimal portfolio at any given point in time can be thought of as roughly a weighted average of the conditionally optimal portfolios, and the weights will be a function of the time horizon and transition probabilities.

[5] There is some confusion in the literature about terminology. Ang & Bekaert (2004) speak of unobserved states, inferred from the data. However, the inferred state is then treated as known and deterministic once identified (which is a by-product of the estimation of parameters).

Ang & Bekaert note that, compared with the setting where the current regime is always treated as unobserved and uncertain, a regime change will have a relatively larger impact on the optimal portfolio. This becomes clearer if one imagines a situation where there is great uncertainty about the prevailing regime: intuition would indicate that the optimal portfolio will be close to evenly weighted to all state-optimal portfolios, regardless of the current state.

They estimate the regime-shifting model with the Bayesian filtering algorithm of Hamilton (1989). Although regime identification is done endogenously with probabilistic methods, once a probability has been assigned that each period belongs to a specific state, the prevailing state is deterministically identified as the state with the highest probability, as if known with certainty. They allow for models where the regimes in the different countries correspond exactly, or alternatively, where regimes in the three countries are not in sync with each other[6], but ultimately find that there is a high degree of correspondence between regimes in each country. The two states were noted to correspond to "normal" periods and periods of both high volatility and high correlations.

In the very specific problem setting of finding the optimal allocation over time that allocates to United Kingdom (UK), German, and US equities, assuming a two-state first-order Markov chain based on monthly data from 1970 to 1997, Ang & Bekaert ultimately conclude that while regime switching satisfactorily generated the asymmetric correlations observed in equity returns, the increased correlations during times of distress do not completely remove the benefits of diversification, and hence they could not explain the "home bias puzzle". Ang & Bekaert also find that the cost of ignoring regimes is only significant if a risk-free asset is available to the investor, and that this cost is of the order of excluding non-local equities from the portfolio.

Ang & Bekaert (2004) built on their earlier paper, expanding the research to include equities, bonds and cash. This study appears to be one of the first involving equities, bonds and cash to refer to the existence of non-linearities and time dependencies in the distributions of asset class returns as patterns that can be exploited via changing asset allocation, as opposed to as stumbling blocks for the traditional MVO theory. The inclusion of fixed income assets is motivated by their observation that interest rates tend to persist and have low volatility when rates are low, while the opposite is observed when rates are high.

First the authors show how a simple Capital Asset Pricing Model (CAPM) framework with the Morgan Stanley Capital International (MSCI) World return as the market and several individual countries as the constituents of the market portfolio can give rise to a wide variety of distributions simply by the introduction of a two-state, first-order Markov chain with constant transition probabilities. Expected returns for the next period with uncertain regimes are simply weighted averages of expected returns with the weights being the probabilities of transitioning to the given state. Expected variance for

[6] For three countries, the latter, more general model would have 2x2x2=8 possible regimes in total and hence 8x8 = 64 transition probabilities.

the next period is a similar weighted average, but due to generally different means in each state, there is an additional "jump" component.

In the second case a "Market-Timing Model" is specified for a US-based investor faced with a choice between equities, bonds and cash.

The cash rate is modelled as a simple autoregressive process, but the constant term and the autoregressive parameter depend on the regime. Equities and bonds have constant expected returns across regimes but the covariance matrix of the three asset classes' error terms depend on the prevailing regime.

They allow the transition probabilities to vary over time, depending on the interest rate, invoking the notion that interest rates affect the probability of transitioning or not between the two regimes, which again correspond to "normal" conditions and "bad" conditions. The transition probabilities are based on simple logit functions of the interest rate. According to statistical tests, the hypothesis that the transition probabilities are constant is strongly rejected.

The long run probability of being in the normal regime implied by the model is 0.70. Although the study is intended to merely illustrate the potential of regime switching models, out-of-sample testing showed their model, which is updated on a monthly basis for incoming data, performs well on a return and risk-adjusted return basis on monthly data from 1985 to 2000. Due to the causal relationship assigned to interest rates and regimes in their model, much of the discussion on the optimal allocations hinges on the level of interest rates. For example, in the "good" regime, when interest rates are low, the investor borrows at the risk-free rate and invests most of the proceeds in equities, as, according to their model, the good regime is likely to persist when interest are low. As interest rates rise, equities become less attractive as the probability of switching to the high volatility regime increases, and the optimal equity allocation is lower. In their problem setting it is evident that the cash allocation is the safe-haven against volatility, significantly more so than bonds.

Honda (2003) is independent of Ang & Bekaert, but refers heavily to the work of Merton (1969) and the concepts of a myopic portfolio and a hedging portfolio. This work, which is entirely theoretical (no attempt is made to fit to real word data), addressed regimes in the context of optimal portfolio choice as well as optimal consumption[7] of risk averse investors with different investment horizons and with "power utility". Regimes changes are again modelled as Markov chains, but in contrast to Ang & Bekaert (2002), regimes are not assumed to be observable and the Markov chain is defined in continuous time. However, Honda's setting is less general in the sense that only the *means* of risky assets depend on this unobservable economic regime. Further, he narrows his research to the case of only two regimes, presumably purely for the sake of simplicity, as no argument is made in favour of a two-regime model. Regimes are estimated endogenously based on all past and present available asset prices, and at any point in time there is a probability of the regime being in either state. Transition times are assumed to follow the same exponential distribution (and are

[7] The rate at which an investor should draw down from his investment at different points in time.

therefore memoryless) regardless of the state. They compute solution sets to their theoretical setting with the help of Monte-Carlo simulation.

Guidolin & Timmerman (2005) attempt to find the multiple-period optimal portfolio holdings allowing periodic rebalancing over time for investors with horizons of varying length who are not concerned with only the mean and variance of terminal wealth, but also the skewness and kurtosis, in a framework that incorporates *unobservable* regimes. Regimes, current or future, are not assumed known with certainty, or assumed to be in the most likely state: instead, there is a probability of being in any of the given states at any point in time. The cost of this more realistic setting is that Monte-Carlo simulation is required to find approximately optimal portfolios.

The analysis is limited to the simple setting of two assets: US stocks and a risk-free asset. Similar to many other works, they find evidence for two regimes in US stocks that were found to roughly correspond to a bear state with high volatility and low mean returns, and a bull state with high mean returns and low volatility. The number of regimes was endogenously inferred from monthly returns of the New York Stock Exchange (NYSE) from 1952 to 1999 using the Schwarz Information Criterion (discussed at length in the context of financial returns by Bossaerts & Hillion, 1999).

Both states reportedly exhibit high persistence, but the bull state (0.95) is more persistent than the bear state (0.8). As a result the average length of a bear state is five months versus 20 months for a bull state. The higher the probability of being in a bull state, the higher is the allocation to equities. When in the bull regime, the amount of equity in the optimal portfolio is decreasing with the investment horizon (as the probability of eventually switching to the bear state increases). Conversely, when in the bear state, the allocation to equities is upward sloping over time horizon, as the probability of eventually switching to a bull state increases.

Both expected returns and covariances are allowed to be state dependent. As in Ang & Bekaert (2002), conditional on the state, the return distribution is assumed to be normal. However, since they do not assume a state is ever known with certainty, the return distribution is always a mixture of normal distributions. Approximately optimal portfolios over time are found with an iterative, Monte-Carlo simulation based algorithm that reduces the problem to several buy-and-hold problems.

Sa-Aadu, Shilling & Tiwari (2005) apply the framework of Hamilton (1989) to the problem of asset allocation for a US investor who has to choose between local equities (large and small caps separately), global equities, real estate investment trusts (REITs), commodities and treasury bonds. They use a two-state Markov chain and find that the estimation procedure of Hamilton (which automatically identifies each period as belonging to a regime) results in regimes that roughly correspond to higher mean and less volatile returns and lower mean and more volatile returns, i.e. a “good” state and a “bad” state. They relate these states informally to the “business cycle”. Good and bad states are noted to be persistent, with a 79% probability of remaining in the good state and a 61% probability of remaining in the bad state from one month to the next.

Celikyurt & Ozekici (2007) restrict themselves to a purely theoretical setting with the aim of finding the evolution of the optimal portfolios over multiple discrete time periods for an investor with quadratic utility in the terminal value of his wealth. They assume that the state of the Markov chain is observable. The means, variances and covariances of the hypothetical assets are also assumed to be known during each state. They solve a purely hypothetical example problem with two states and two assets (a risky and a risk-free asset) with dynamic programming methods.

Guidolin & Timmerman (2007) extend their earlier work to the setting of equities (large and small cap separately), bonds and cash, using monthly data from 1954 to 1999. In this setting, they find evidence that four regimes are needed to capture the joint distribution (as opposed to the two regimes of their 2005 study). Hamilton's estimation method of maximising the likelihood function is adapted to employ Monte-Carlo methods to enable states to remain unobserved and probabilistic. As before, information criteria and several specification tests are considered to select the number of states, which varied between one and six. None of these six models passed all tests, but a four-state scheme (with state dependent means as well as variances) was adjudged to be the most parsimonious model that adequately captured the distribution of all asset classes.

The regimes were found to correspond to: (1) a crash state (negative mean returns and high volatility); (2) low growth (low volatility and modest positive returns for all asset classes); (3) a bull state (high returns on stocks, negative returns for bonds); and (4) a recovery state (high returns and high volatility). Correlations were also noted to vary between regimes, with the correlation between large stocks and bonds as low as -0.4 in the crash state and as high as 0.37 in the recovery state.

States 1 and 4, the crash (average duration two months) and recovery states (average duration three months) were far less persistent than the other two states (average durations of seven and eight months), and found to roughly correspond to the NBER recession periods. The transition probabilities of the Markov chain imply that the long-term proportion of time spent in each regime are for state 1 to 4, respectively, 9%, 40%, 28% and 23%.

The persistence of states and transition probabilities between states are noted to result in the initial state (i.e. current state) probabilities being important determinants of optimal allocations. The authors discuss the intuitively pleasing relationships between the current perceived regime and the aggressiveness of optimal portfolios. A similar relationship as in the authors' earlier work of the probabilities of the current state and time horizon is reported. The effects of allowing for states to be uncertain as opposed to observed is adjudged to be significant. The effect of time horizon is, however, greatly diminished as the rebalancing frequency is increased.

Tu (2010) considers the question of optimal allocation in the presence of regimes in a Bayesian framework on a monthly dataset of stock returns from 1963 to 2006. Returns are assumed to follow a joint normal distribution conditional on the regime. The author makes use of the Bayesian framework's inherent ability to explicitly incorporate model and parameter uncertainty. The model assumed in this case is the Fama & French three-factor model (1993). One motivation for using this model is that the Fama & French study involves 28 risky assets, allowing Tu to showcase that the Bayesian approach can cope with a large number of assets. The computational strain in a setting with such a large number of asset classes on the non-Bayesian approaches of previous research would, according to Tu, render those methods impractical.

Tu then settles on two distinct regimes, apparently largely because a higher number of regimes result in some regimes having very few months of data. The two regimes are described as a bull regime with high returns and lower volatility and a bear regime with lower returns and higher volatility. Correlations are also found to be regime dependent.

Finally Tu finds that the economic value of including regimes is positive, even after accounting for model and parameter uncertainty.

Turner & Han (2009) divide historical periods into different clusters on the basis of 27 economic indicators (i.e. exogenously with respect to asset returns). The implicit assumption is that the state at any point in time is known with certainty. Standard MVO is then performed separately for each cluster for a set of equity sector indices of US stocks between 1980 and 2009.

They first reduce the dimensionality of the problem with principal components analysis to two dimensions, with the dimensions interpreted as corresponding to economic states: The first principal component is found to correspond consistently with certain positive and negative economic indicators, while the second appears to correspond to a large degree with only inflation data. The number of clusters is determined by testing at each point in time the hypothesis that the prevailing economic state is in an existing cluster, and creating a new cluster if it is not.

They found that new observations very frequently do not fall within an existing cluster, necessitating an apparently arbitrarily large number of clusters as the period in consideration increases in length. This prompts the authors to use all observations at any point in time, weighing data according to the distance from the current regime in the principal component analysis (PCA) space. Both the discrete and continuous methods outperformed the benchmark in terms of both return and variance.

Ang & Timmerman (2012) provide an overview of regime switching models in various contexts: equity returns, interest rates, exchange rates and asset allocation. They discuss the motivations for employing these models and the phenomena these models are able to reproduce, such as skewness, kurtosis, time-varying correlations, transitory shocks, and the non-monotonic relationship between expected return and variance.

The approach of **Kritzman, Page & Turkington (2012)** employs both exogenous and endogenous time series to identify regimes in a Markov-switching model. They arbitrarily fit two-state Markov processes, separately to each of their three regime identifying time series, namely market turbulence, inflation and economic growth. Market turbulence is defined as the divergence (as measured by the Mahalanobis distance) of asset returns from the sample mean and sample covariance over the full period, and is the endogenous variable used in identifying regimes. They find that in all three cases the two regimes correspond respectively with a normal state, and an “event” state. The event state in all three cases experiences lower returns with higher volatilities. Regimes are fitted via likelihood maximisation with an expectation maximisation algorithm. In out-of-sample testing, they find that a simple dynamic asset allocation strategy that tilts a portfolio depending on the identified regimes, outperforms a static allocation, with the breakeven transaction cost being 133 bps. The authors relate the outperformance to the ability to predict higher levels of volatility to persistent regimes, rather than the effects on mean returns. Due to the high persistence of regimes, turnover was limited to less than 150% per year.

Bae, Kim & Mulvey (2014) apply Hidden Markov Models to the problem-setting of allocating to three asset classes, namely Standard & Poor’s 500 Index (S&P 500), US bonds and a commodities index, employing daily data over the period 1980 to 2012. Conditional on the regime, the three asset classes are assumed to follow a joint normal distribution. The Baum-Welch expectation maximisation algorithm (see Fraser, 2008) is used to estimate the parameters.

To select the number of states, the authors use the Bayesian Information Criteria along with heuristic arguments based on inspecting the transition matrices of respectively 2, 3 and 4 state Markov models, ultimately settling on a four-state model. State 4 is the “worst” state and involves high volatility for all three assets, low to negative mean equity and commodity returns, and high bond returns. State 1 is the good state, and roughly the inverse of state 4. States 2 and 3 are transition states between states 1 and 4, and exhibit properties somewhere in-between. The correlations are also seen to be highly regime dependent. All states have persistence of greater than 0.6.

The fitted Hidden Markov Model is then used to simulate returns sequentially over time, at each step employing a longer set of data to incorporate new observations. Employing dynamic programming methods, they find optimal portfolios over time, each maximising one-period utility, given the latest fitted model. The optimal portfolio takes into account parameter uncertainty by an averaging of the optimal portfolio for each simulated outcome, over many simulations. They find that the model outperforms several benchmark portfolios, especially adding value during crash periods by lowering the risk in the portfolio at these times.

Nystrup et al. (2015) set out to investigate whether regime-based asset allocation can outperform static portfolios in the absence of the ability to predict regimes. They endogenously fit two regimes (once again a low volatility and high volatility categorisation) on daily data of global equities (MSCI All Countries) and global government bonds (JPMorgan Global Government Bond Index). Instead of attempting to predict future regimes, the current regime is identified with Hidden Markov Model methods on daily data in real time, and their simple dynamic, regime-based allocation strategies rely on the persistence of regimes to add value relative to a static portfolio. For example, one strategy switches between 100% equities and 100% bonds if there is greater than 95% confidence that a regime change has taken place on a specific day. The weight given to older data decays exponentially to account for gradual changes in the parameters of the data generating processes.

They find the stocks versus bonds strategy has a better Sharpe ratio than a static portfolio if the transaction costs do not exceed 239 bps, and also improved maximum drawdowns. The authors conclude that if the persistence of regimes noted in the past continues in future, regime-based asset allocation can add value regardless of whether regimes can be predicted.

3. KEY CONCEPTS AND DEFINITIONS

In this chapter the distributions of asset returns and the assumptions commonly made in practice and in the literature are briefly discussed in Section 3.1. This will form the backdrop against which some key concepts employed in this research, among which several different measures of risk, will be defined and discussed in Section 3.2.

3.1. The distribution of returns

Asset prices are often assumed to follow geometric Brownian motion with constant drift and constant volatility:[8]

$$dP_t = \mu^* P_t dt + \sigma P_t dz_t \tag{1}$$

where dz_t is the Wiener process (standard Brownian motion).

It follows via Ito's Lemma that (see Tsay, 2005:228)

$$dlnP_t = \left(\mu^* - \frac{\sigma^2}{2}\right) dt + \sigma dz_t \tag{2}$$

Integrating on both sides over t to $t + T$ yields

$$lnP_{t+T} - lnP_t = \left((\mu^* - \frac{\sigma^2}{2})(T + t - t) + \sigma(z_{t+T} - z_t)\right) \tag{3}$$

From the properties of the Wiener process, it follows that

$$lnP_{t+T} - lnP_t \sim normal\left(T(\mu^* - \frac{\sigma^2}{2}), T\sigma^2\right) \tag{4}$$

Thus, if we define $\mu \equiv \mu^* - \frac{\sigma^2}{2}$, then lnP_t follows a generalised Wiener process with drift μ and volatility σ. It also follows that if we define $R_{t,t+T}$ as the (non-annualised) return over the time interval t to $t + T$, that

---

[8] See Merton (1969); Merton (1973); and Black & Scholes (1973) for some of the earliest allusions to this model in the financial context.

$$R_{t,t+T} \equiv ln\frac{P_{t+T}}{P_t} = lnP_{t+T} - lnP_t \sim normal(T\mu, T\sigma^2) \tag{5}$$

and

$$lnP_{t+T}|lnP_t \sim normal(lnP_t + T\mu, T\sigma^2) \tag{6}$$

where $lnP_{t+T}|lnP_t$ denotes the distribution of lnP_{t+T} conditional on lnP_t.

If there is more than one asset under consideration, (5) can be generalised to the multivariate case:

$$\boldsymbol{R_{t,t+T}} \sim normal(T\boldsymbol{\mu}, T\boldsymbol{\Sigma}) \tag{7}$$

where, assuming there are k assets, $\boldsymbol{R_{t,t+T}}$ is a $kx1$ random vector with the k asset returns for the period t to $t+T$, $\boldsymbol{\mu}$ is the $kx1$ vector of the true means of the distributions of asset returns, and $\boldsymbol{\Sigma}$ is the kxk true covariance matrix of the distributions of asset returns.

From the definition of the Wiener process, successive infinitesimal increments are independent, and the returns achieved over any two non-overlapping periods are clearly independent from each other.

It is clear from the above results that under the assumption of geometric Brownian motion with constant drift and constant volatility, the following statements about the distribution of returns hold true:

1. Returns have a stationary distribution for any given periodicity;
2. Returns are normally distributed and therefore the first two moments (and co-moments if there is more than one asset), mean and variance (and covariance), fully capture returns distributions; and
3. The variance of returns over any given periodicity is a constant factor of the variance of returns over any other periodicity.

Note that the last point does not actually require the assumption of geometric Brownian motion with constant drift and volatility, or even an assumption of normality. It is only necessary that returns have a stationary distribution and are independent of each other across time (in other words that periodic returns are i.i.d.). To see this, consider the following for a single asset's return between times t and $t+T$:

$$\sigma^2_{t,t+T} \equiv var(R_{t,t+T}) = E[R_{t,t+T} - E(R_{t,t+T})]^2$$
$$= E\left[\sum_{i=t}^{t+T-1} R_{i,i+1} - E(\sum_{i=t}^{t+T-1} R_{i,i+1})\right]^2$$

$$= E\left[\sum_{i=t}^{t+T-1} R_{i,i+1} - \sum_{i=t}^{t+T-1} E(R_{i,i+1})\right]^2$$

$$= E\left[\sum_{i=t}^{t+T-1} \left(R_{i,i+1} - E(R_{i,i+1})\right)\right]^2$$

$$= E\left[\sum_{i=t}^{t+T-1} \left(R_{i,i+1} - E(R_{i,i+1})\right)^2\right]$$

$$+ E\left[\sum_{i=t}^{t+T-1} \sum_{\substack{j=t \\ i\neq j}}^{t+T-1} \left(R_{i,i+1} - E(R_{i,i+1})\right)\left(R_{j,j+1} - E(R_{j,j+1})\right)\right]$$

$$= \sum_{i=t}^{t+T-1} E\left(R_{i,i+1} - E(R_{i,i+1})\right)^2$$

$$+ \sum_{i=t}^{t+T-1} \sum_{\substack{j=t \\ i\neq j}}^{t+T-1} E\left(R_{i,i+1} - E(R_{i,i+1})\right)\left(R_{j,j+1} - E(R_{j,j+1})\right)$$

$$= \sum_{i=t}^{t+T-1} var(R_{i,i+1}) \; + \; 2 \sum_{i=t}^{t+T-1} \sum_{\substack{j=t \\ j<i}}^{t+T-1} cov(R_{i,i+1}, R_{j,j+1}) \tag{8}$$

If non-overlapping monthly returns are identically distributed, the first expression reduces to $T\sigma_1^2$. If returns are also assumed to be uncorrelated, the summation of covariances is zero. So we have shown that a sufficient condition for the relationship

$$\sigma_T^2 \equiv \sigma_{t,t+T}^2 = T\sigma_1^2 \tag{9}$$

to hold, is for consecutive returns to be independently and identically distributed (of course, for the mean and variance to fully describe the distribution of returns, we still *do* require returns to be normally distributed). Most studies and industry research on the optimal asset allocation for the medium to long term estimate the variance of returns with the sample variance based on monthly data, implicitly assuming that monthly returns are i.i.d. and normally distributed.

However, if consecutive returns are generally positively correlated (in violation of the independence of the Wiener process and thus the constant geometric Brownian motion model), the covariance terms will be positive and we can see from (8) that

$$\sigma_T^2 > \; T\sigma_1^2 \tag{10}$$

Conversely, if consecutive returns are generally negatively correlated, the covariance terms will be negative and

$$\sigma_T^2 < T\sigma_1^2 \tag{11}$$

If the covariance terms are positive for returns that are closer to each other in time, and negative for returns that are further away from each other in time, the relationship between σ_T^2 and $T\sigma_1^2$ will be dependent on T:

$$\sigma_T^2 < T\sigma_1^2 \; for \; T > T_0 \tag{12}$$

and

$$\sigma_T^2 > T\sigma_1^2 \; for \; T \leq T_0 \tag{13}$$

for some positive constant T_0.

The literature on equity market returns refers to the phenomena *short-term momentum* for returns for T around 12 months and *long-term mean reversion* for returns of around 36 to 60 months (Poterba & Summers, 1986; Fama & French, 1988; and Fama & French, 2012). Thus employing the sample variance of monthly returns should tend to underestimate the variance of $R_{t,t+T}$ if T is around 12 months and overestimate it if T is between 36 to 60 months.

It is similarly possible, in principle, for a relationship to exist between the returns of two *different* assets across asynchronous periods of time. For example, the covariance between the returns of two different assets, denoted by $R_{t,t+T}$ and $S_{t,t+T}$, can similarly be shown to be

$$cov(R_{t,t+T}, S_{t,t+T}) = \sum_{i=t}^{t+T-1} cov(R_{i,i+1}, S_{i,i+1}) + 2\sum_{i=t}^{t+T-1} \sum_{\substack{j=t \\ j<i}}^{t+T-1} cov(R_{i,i+1}, S_{j,j+1}) \tag{14}$$

If the terms after the plus sign are non-zero, the covariance between two assets' returns may be underestimated or overestimated by sample estimates based on monthly returns.

A potential example of such a scenario is a delayed effect of low interest rates on equity returns. Lowering interest rates may have an immediate buoying effect on equity returns due to changing market expectations for the economy. It may, however, also have the effect of boosting the economy, which in turn may have positive second-order effects on equity markets. However, some period of time greater than a month may have to elapse for this sequence of events to come to bear. In other words, a low interest right now may have an effect on equity markets in say 12 months' time, for example.

Similar expressions to the right-hand side of (14) exist for the higher order moments and co-moments of multi-period asset returns would similarly show how these quantities are affected by non-independence between asset returns across time.

An example of how the geometric Brownian motion model can be adjusted to allow for autoregression that results in short-term momentum and long-term mean reversion is given by Koijen & Rodriguez (2009:5):

$$dP_t = (\phi M_t + (1-\phi)\mu_t)P_t dt + P_t \boldsymbol{\sigma'}_{\boldsymbol{S}} \boldsymbol{dz_t} \tag{15}$$

where both $\boldsymbol{\sigma'}_{\boldsymbol{S}}$ and $\boldsymbol{z_t}$ are two-dimensional vectors that allow for independent return and dividend yield shocks. M_t is a weighted sum of past returns:

$$M_t = \int_0^t e^{-(t-u)} \frac{dP_t}{P_t} \tag{16}$$

with $e^{-(t-u)}$ representing the time-weighting scheme.

Observed real world asset prices and returns are known to violate constant geometric Brownian motion and the normal distribution for returns in many other ways as well. To mention but a few:

- Skewness and kurtosis in various asset classes (Schwert, 1989; and Cont, 2001);
- Co-skewness and co-kurtosis between asset classes (Guidolin & Timmerman, 2008);
- Persistent volatility (Schwert, 1989; and Cont, 2001);
- Time-varying correlations (especially an increase during market downturns) (Longin & Solnik, 2001; and Ang & Chen, 2002); and
- Autoregressive characteristics: Short-term (approximately 12 months) momentum and long-term mean reversion (36 to 60 months), as already mentioned above.

Most of these violations can be captured by allowing the drift and volatility of the geometric Brownian motion process to vary across time. In other words:

$$dP_t = \mu_t^* P_t dt + \sigma_t P_t dz_t \tag{17}$$

where μ_t^* and σ_t are some functions of time. In general, the distributions of returns based on this class of processes are of course not stationary, which may render simple point and density estimates based on all observed returns meaningless, or not strictly relevant to the distributions of future returns.

If the time-dependent drift and volatility can be defined in such a manner that stationarity is retrieved, or retrieved in a restricted sense, general normality may be lost but estimates of moments based on all historical returns may still be relevant (if limited in their ability to describe the full return generating process). In such cases the distribution of returns may no longer be fully described by the first two moments and mean-variance optimisation may not be appropriate to investors with utility functions that are not indifferent to higher order moments.

One approach that potentially retrieves stationarity in an unconditional sense is to hypothesise the existence of "regimes" or "states", conditional on which stationarity or even normality is assumed to be restored.

For example, in the context of geometric Brownian motion, *conditional on the state at time* t, constant drift and constant volatility may be assumed to be retrieved, and conditional returns are once again normally distributed:

$$R_{t,t+1}|S(t) = i \sim normal(\mu_i, \sigma_i^2) \tag{18}$$

where $S(t)$ denotes the state at time t and μ_i and σ_i^2 are respectively the conditional mean and variance under state i.

If the probability of being in any given state at any specific point in time is invariant across time and not dependent on the prevailing state, the resultant unconditional distribution of returns is a mixture of normals where the weight on each normal density is this constant probability:

$$f_{R_{t,t+1}}(x) = \sum_{i=1}^{N_s} p(i)\phi(x, \mu_i, \sigma_i^2) \tag{19}$$

where $p(i) \equiv P(S(t) = i)$, $\phi(\mu_i, \sigma_i^2)$ is the normal density[9], and N_s is the number of states.

The ability of this model to capture persistent changes in drift and variance is obvious. However, various other phenomena can be explained by this model. To illustrate the effect on moments we show below the first four moments for $N_s = 2$ and $p \equiv p(1)$, (Ang & Timmerman, 2012:5):

$$ER_{t,t+1} = p\mu_1 + (1-p)\mu_2 \tag{20}$$

$$var(R_{t,t+1}) = p(1-p)(\mu_1 - \mu_2)^2 + p\sigma_1^2 + (1-p)\sigma_2^2 \tag{21}$$

$$skew(R_{t,t+1}) = p(1-p)(\mu_1 - \mu_2)[(1-2p)(\mu_1 - \mu_2)^2 + 3(\sigma_1^2 - \sigma_2^2)] \tag{22}$$

[9] Note that one need not assume the conditional distributions are normal. "Markov-switching models" can be defined relative to any family of distributions, or, for example, any generalised autoregressive process. However, it is typical to hypothesise that conditional normality is restored within states, and this is consistent with constant geometric Brownian motion within states.

$$kurt\left(R_{t,t+1}\right)=p(1-p)(\mu_1-\mu_2)^2[((1-p)^3+p^3)(\mu_1-\mu_2)^2+6p\sigma_2^2+6(1-p)\sigma_1^2)]+3p\sigma_1^4+3(1-p)\sigma_2^4 \tag{23}$$

It is evident how even this simple model can generate skewness and kurtosis. Notice that there has to be a difference in means for skewness to exist. Variance is not simply a weighted average of the two variances, but is increased by the difference in mean: the possibility of jumping from one mean to another increases the overall variability (hence variance said to have a *jump component*). In fact, the difference in mean impacts all four of these moments.

The above model can be further generalised by allowing states to change according to a discrete Markov chain where the probabilities depend on earlier states. Typically a first-order Markov chain is employed, i.e. the probability of being in a certain state at a specific point in time depends only on the previous state. For example, the probability of transitioning from state i at time t to state j at time $t+1$ is:

$$p^{(t,t+1)}(i,j) \equiv P(S(t+1)=j|S(t)=i) \tag{24}$$

In this scenario, the unconditional distribution of returns is also a mixture of normals but here the weight on each normal density depends on the last known state. For example, if the state at time t is known and state changes occur only at discrete time intervals, the resultant probability density function (pdf) for the distribution of returns for the interval $t+1$ to $t+2$ is

$$f_{R_{t+1,t+2}}(x \mid S(t)=i\,)=\sum_{j=1}^{N_s} p^{(t,t+1)}(i,j)\phi(x,\mu_j,\sigma_j^2) \tag{25}$$

The density for any number of periods s into the future can be stated with reference to the relevant s-period forward probabilities (which in turn can be stated in terms of the one-period forward probabilities).

In general, if the N_sxN_s state-changing probability matrix does not change over time (i.e. is "homogenous") and under certain regulatory conditions (e.g. ergodicity), the state-changing mechanism can be stationary. As we can see, in this scenario, if the prevailing state is not known with certainty, the unconditional return-generating process is non-normal but stationary, and its moments exist and can be estimated with historical returns. However, the sample moments based on all data are in such a scenario at best limited in their ability to describe the return behaviour, as they completely ignore the existence of regimes.

As discussed in the literature review, states are generally found to persist from one to month to the next with a probability of typically between 0.6 and 0.95, which means they persist for an expected time of about 2.5 to 20 months.

One of the biggest challenges in defining these Markov-switching models is deciding on the number of states or regimes (Ang & Timmerman, 2012:4). Comparing the fit of say a two-state model to that of a three-state model in the presence of unknown structural parameters (e.g. mean and variance in the case of normal conditionals) for each state, as well as all the transition probabilities involved, poses a significant technical challenge. As a result, some researchers (Ang & Timmerman, 2012:4) argue that the number of regimes should be based on fundamental (e.g. economic) arguments. Others argue for the use of statistical tests (Hamilton, 1996) or information criteria (Psaradakis & Spagnolo, 2003) to choose the regimes. Often the choice is arbitrarily made, and in those cases it is usually two regimes, a "good" regime and a "bad" regime.

A further challenge faced by these methods is that there is no guarantee that the conditional return distributions truly conform to the assumed distribution, e.g. the normal distribution in our examples above. This issue is typically ignored in the regimes literature.

The discussion above and the implications can be summarised as follows:

1. Observed unconditional asset returns tend not to be distributed identically and independently normal, and at least all of the following phenomena have all been observed in actual asset returns: skewness, kurtosis, momentum, mean reversion, persistent volatility. Simple point estimates of mean and variance therefore do not in general fully capture the distributions of returns.

   Further, variance for any arbitrary period cannot be estimated with data with any arbitrary periodicity – variance, skewness and kurtosis (and their co-equivalents) may be dependent on the periodicity. Thus, the very common practice in the industry of estimating risk purely with the sample covariances matrix of monthly returns ignoring the regime-dependence of distributions may lead to biased estimates of the entries in the covariance matrix for longer investment horizons.

2. A popular alternative approach is to model the returns generating process with a Markov-switching model. Even then, conditional returns may not conform to the assumed distribution. In the case of the normal density assumption, skewness and kurtosis may still be present in the observed returns, even after dividing returns into the different states.

3. In the presence of regimes, the unconditional distribution is a mixture of the assumed conditional densities, and there will in general be both skewness and kurtosis present in the unconditional distributions (even if conditional distributions are assumed to be normal).

4. Misspecification of the number of regimes or the conditional densities and the misclassification of a period to a regime are possible sources of model error.

3.2. Definitions

3.2.1. The required real return and investment horizon

The **required real return** is the annualised real return threshold that must be achieved, often over an explicitly defined period of time, the **investment horizon**. Thus the expected return of the chosen portfolio should at the very least be in excess of the required real return, and risk should be measured in a manner that takes into account the fact that the required return should be delivered over the investment horizon.

Define ${R'}_{(j)}^{t,t+T}$ as the ordinary (non-logarithmic) non-annualised return on asset class j between times t and $t+T$.[10] Define $R'^{(p)}_{t,t+T}$ as the ordinary (non-logarithmic) non-annualised return between times t and $t+T$ on portfolio p, defined by weights $w_1^{(p)}, w_2^{(p)}, \dots, w_k^{(p)}$ to the asset classes .

We can express $R'^{(p)}_{t,t+T}$ in terms of monthly returns as follows:

$$R'^{(p)}_{t,t+T} = \prod_{i=t}^{t+T-1} (1 + R'^{(p)}_{i,i+1}) - 1 \tag{26}$$

We assume a full monthly rebalancing. Hence we can express $R'^{(p)}_{t,t+T}$ in terms of our raw data, namely monthly (non-logarithmic and non-annualised) asset class returns, as follows:

$$R'^{(p)}_{t,t+T} = \prod_{i=t}^{t+T-1} \left(1 + \sum_{j=1}^{k} w_j^{(p)} {R'}_{(j)}^{i,i+1}\right) - 1 \tag{27}$$

We can now define $R^{(p)}_{t,t+T}$ as the logarithmic non-annualised return on portfolio p and write it in terms of the monthly ordinary asset class returns and the portfolio weights:

$$\begin{aligned} R^{(p)}_{t,t+T} &= \ln\left(1 + R'^{(p)}_{t,t+T}\right) \\ &= \ln\left[\prod_{i=t}^{t+T-1} \left(1 + \sum_{j=1}^{k} w_j^{(p)} {R'}_{(j)}^{i,i+1}\right)\right] \end{aligned} \tag{28}$$

[10] As a general rule, upper case letters will be used to denote random variables, and lower case to denote their observed outcomes. One exception to this rule is T for time horizon. Not also that in order to clearly distinguish between the return on a *portfolio* and the return on an *asset*, time is denoted by the subscript for portfolios and by the superscript for assets.

We analogously define $I_{t,t+T}$ as the non-annualised logarithmic inflation rate between times t and $t+T$, based on the South African Consumer Price Index (CPI) basket of goods and services.

Lastly, we define $U_{t,t+T}^{(p)}$ as the annualised *real* (logarithmic) return of portfolio p over the period t to $t+T$ (measured in months):

$$U_{t,t+T}^{(p)} = \frac{12}{T}(R_{t,t+T}^{(p)} - I_{t,t+T}) \tag{29}$$

3.2.2. The strategic and tactical asset allocation

In the context of financial planning, the required real return may refer to the outcome of a detailed future cash flows assessment of an individual or institution. Typically this assessment will be a simple time value of money exercise, and the required real return that emanates for this process is a constant real return that balances inflows and outflows. The individual or institution's assets are then managed with the objective of meeting this real return. Often the key consideration in this setting is the *asset allocation* to a handful of *asset classes*, typically including local equities, bonds, property and cash and their global equivalents. Typically there are two broad categories of asset allocations: **the strategic asset allocation** – the long-term benchmark allocation – and the **tactical asset allocation** – the targeted asset allocation at any point in time.

The strategic asset allocation (SAA) is typically long-term in nature, and serves as a benchmark with which to compare the actual performance of the fund or the tactical asset allocation (TAA). The SAA is usually the asset allocation that is deemed best positioned to achieve a fund's objective (which is often a real return) in the absence of tactical asset allocation. As such it is often stated on marketing material and the profile documentation of products. Underlying the SAA are typically assumptions about the future expected returns, risk, and the relationships between the asset classes. These assumptions are usually informed by the long-term historical returns of asset classes, quantitative models, consensus opinions in the market and the long-term views of the asset manager.

The tactical asset allocation is the asset allocation that is being targeted by the fund manager at any point in time. It takes into account the prevailing economic and market conditions and the fund manager's views on all relevant factors and represents the allocation best placed, in the short term, to meet the objectives of the fund, or to outperform the SAA, or that is optimally positioned from a risk-return perspective. In statistical terms the TAA can be seen as the optimal allocation conditional on the current state (where the state may be measured by exogenous or endogenous variables).

The exact justification given by a fund manager for veering from the SAA is highly dependent on the current market, economic and political circumstances, as well as the investment management style and philosophy of the manager. The performance of the tactical asset allocation is usually measured against the strategic asset allocation and the objectives of the fund.

An important question in the context sketched above is: What are the strategic and tactical asset allocations that are optimal in targeting the required real return over the investment horizon? The answer depends on what is meant by optimal. The typical quantitative approach is to simply find the portfolio with the minimum standard deviation (where standard deviation is estimated by the sample covariance matrix based on monthly returns) that meets the required real return, i.e. mean-variance optimisation. The expected returns and the covariance matrix may be purely based on history, or be the outcome of a quantitative or qualitative analysis. However, if one thinks about it, under the assumptions of this approach, this optimal portfolio should have a roughly 50/50 chance of meeting the required return.

Would a more aggressive portfolio that targets a higher real return not have a higher probability of achieving the required real return over longer investment horizons? Also, would this probability not increase the longer the investment horizon, thereby making it more and more attractive as the investment horizon increases in length? On the surface of things that does seem to make sense. However, at the same time, this more aggressive portfolio is likely to incur higher risk in the traditional, mean-variance sense. It is therefore also more likely to severely underperform the required real return.

It becomes clear that the optimal portfolio is dependent on how risk is defined. Further, if risk is defined as underperforming the required real return over the recommended investment horizon, a penalty function on this underperformance has to be specified. This fact and the fact that the distribution of return depends on periodicity, lead us to consider the various risk criteria introduced in the next section.

3.2.3. Standard deviation of returns

This risk criterion measures the variation of the annualised return over the investment horizon of T months:

We have already defined σ_T^2

$$\sigma_T^2 \equiv var(R_{t,t+T}) \tag{30}$$

and shown that if consecutive returns are i.i.d., then the variance over any period is a constant factor of the variance over any other period, and stating variance over any one periodicity is equivalent to stating it over any other periodicity:

$$\sigma_T^2 = T\sigma_1^2 \tag{31}$$

Under these conditions the most efficient estimator of σ_T^2 is (assuming access to monthly data):

$$T\hat{\sigma}_1^2 = T * \frac{1}{N_1 - 1} \sum_{i=0}^{N_1-1} (r_{i,i+1} - \bar{r}_{1,N_1})^2 \tag{32}$$

where $r_{0,1}, r_{1,2}, \dots, r_{N_1-1,N_1}$ are observed monthly returns, $\bar{r}_{1,N_1}$ is the observed mean monthly return and $\hat{\sigma}_1^2$ is of course the sample variance of monthly returns, which is well known to be an unbiased estimator of σ_1^2 (thus $T\hat{\sigma}_1^2$ is also an unbiased estimator of σ_T^2 under these conditions).

However, as already discussed in Section 3.1, if $R_{t,t+T}$ is a stationary process but not independent across time, the covariance terms may be non-zero, and

$$\sigma_T^2 \neq T\sigma_1^2 \tag{33}$$

If the covariance terms are generally non-zero, then, in general, $T\hat{\sigma}_1^2$ is biased estimator of σ_T^2 **and the estimation of risk, as measured by variance, is dependent on investment horizon**. In this scenario, the natural estimator is the sample variance of T-month returns:

$$\hat{\sigma}_T^2 \equiv \frac{1}{N_T - 1} \sum_{i=0}^{N_T-1} (r_{iT,(i+1)T} - \bar{r}_{T,N_T})^2 \tag{34}$$

where $r_{0,T}, r_{T,2T}, \dots, r_{(N_T-1)T,N_TT}$ are the N_T consecutive (non-overlapping) observed T-month returns, and $\bar{r}_{T,N_T}$ is their mean.

As already mentioned, there is in fact reason to believe that $\sigma_T^2 \neq T\sigma_1^2$ and hence we will distinguish in this work between the two estimates of variance above. We will refer to $T^{1/2}\hat{\sigma}_1$ as the annualised sample monthly standard deviation (or *monthly standard deviation* for short), and $\hat{\sigma}_T^2$ as the sample standard deviation over a T-month investment horizon (or *standard deviation over horizon*, for short)[11].

Lastly, we note that if returns can be assumed to be normally distributed, these statistics (along with the mean returns), fully capture the distributions of returns. If returns are non-normal but stationary, they are still meaningful, but an incomplete description of return distributions for investors who are concerned with skewness and kurtosis.

---

[11] More specifically, our results section will refer to the case where the periodicity is equal to the investment horizon as the “empirical case”. The results in Chapter 5 will in fact examine periodicities that vary between 1 and T months (where T denotes the investment horizon). The same holds for all the risk criteria introduced and discussed below.

All of the remaining risk measures defined below will be defined relative to an explicit investment horizon (that is generally greater than one month). The implicit assumption is that $R_{t,t+T}$ has a stationary distribution, but that it is composed of monthly returns that may exhibit dependencies over time – consistent with the discussion above arguing for the use of $\hat{\sigma}_T^2$ to estimate variance. We will also attempt to account for regimes at a later stage. In the regime cognisant section, it would be more accurate to say that we assume that conditional on the regime, $R_{t,t+T}$ is stationary.

3.2.4. Probability of success

The *probability of success* is defined as the probability of achieving the required real return over the investment horizon:

$$P\left[U_{t,t+T} \geq req\ real\ return\right] \tag{35}$$

Again, there is an implicit assumption that $R_{t,t+T}$ is stationary, though not necessarily normally distributed.

This probability will be estimated by

$$\frac{1}{N_T}\sum_{i=0}^{N_T-1} I(u_{iT,(i+1)T} \geq req\ real\ return) \tag{36}$$

The penalty on underperformance here is clearly binary: the penalty is 1 if there is underperformance and 0 if there is no underperformance. It is irrelevant to this measure exactly *how much* we underperform the required real return.

In contrast to variance, which by definition is only dependent on the first two moments of the distribution of returns, the probability of success is dependent on other attributes of the distribution (and by implication on the higher order moments).

3.2.5. Expected shortfall over investment horizon

This risk measure is defined as the expected difference between the required real return and the achieved real return if there is underperformance, else it is zero.

$$E\left[(req\ real\ return - U_{t,t+T})I(U_{t,t+T} \leq req\ real\ return)\right] \tag{37}$$

This parameter will be estimated by

$$\frac{1}{N_T}\sum_{i=0}^{N_T-1}\left(req\ real\ return - u_{iT,(i+1)T}\right)I\left(u_{iT,(i+1)T} \leq req\ real\ return\right) \qquad (38)$$

As was the case for probability of success, the penalty is zero if there is outperformance of the required real return over the investment horizon period. However, here the penalty for underperformance is linearly increasing in the underperformance (on a per annum basis).

As is the case for probability of success, the expected shortfall cannot be expressed in terms of the first two moments of returns. As a result the optimal expected shortfall portfolio is not necessarily on the mean-variance efficient frontier, and may be dependent on both skewness and kurtosis (and higher-order moments).

3.2.6. Downside deviation over investment horizon

This parameter is very similar to the expected shortfall, with no penalty for outperformance, but here the penalty is quadratic in the underperformance:

$$E\left[\left(req\ return - U_{t,t+T}\right)^2 I(U_{t,t+T} \leq req\ real\ return)\right] \qquad (39)$$

The estimator in this case will be

$$\frac{1}{N_T}\sum_{i=0}^{N_T-1}(req\ real\ return - u_{iT,(i+1)T})I(u_{iT,(i+1)T} \leq req\ real\ return) \qquad (40)$$

In other words, where the average shortfall penalised larger underperformance relatively more severely than probability of success, downside deviation penalises underperformance even more harshly than average shortfall.

Downside deviation is also potentially dependent on the higher-order moments of returns.

3.2.7. Expected maximum drawdown

The maximum drawdown for a portfolio p can be defined as

$$MD_{t,t+T} = \min_{0 \le s \le T} \left[R^{(p)}_{t,t+s} - R^{(p)(max)}_{t,t+s} \right] \quad (41)$$

where $R^{(p)(max)}_{t,t+s}$ is defined as

$$R^{(p)(max)}_{t,t+s} = \max_{0<i \le s} \left[R^{(p)}_{t,t+i} \right] \quad (42)$$

The expected maximum drawdown over period $\{t, t+T\}$ is defined simply as:

$$E[MD_{t,t+T}] \quad (43)$$

This parameter will be estimated by the average maximum drawdown:

$$\overline{MD} = \frac{1}{N_T} \sum_{j=1}^{N_T} md^{(j)}_T \quad (44)$$

where $md^{(j)}_T$ is maximum drawdown recorded in the j^{th} return period of length T.

It is often claimed in the financial planning industry that investors who are withdrawing (disinvesting) amounts from an investment are more averse to large drawdowns than to month-on-month variations in returns, as he or she would not participate fully in the subsequent recovery due to the regular withdrawal of assets. While it is obvious there will be a direct link between monthly standard deviation and maximum drawdown, there may be dependency structures in asset class returns over time that render the optimal monthly standard deviation portfolio and the optimal average maximum drawdown sufficiently distinct from each other to usefully distinguish between the two. It is possible that an investor who is regularly withdrawing from an investment would be better positioned in a portfolio that minimises drawdowns rather than monthly standard deviation.

The maximum drawdown is only related to investment horizon in that a short investment horizon is more likely to cut short a drawdown in progress. However, our results section will show that the maximum drawdown optimal portfolio is essentially independent of investment horizon.

3.3. Discussion of the various risk measures

The monthly standard deviation is perhaps the default risk measure in the investment industry and academic literature. It assumes that monthly returns are identically and independently distributed (or at least i.i.d. within regimes), and hence that month-on-month variation fully captures risk.

As already discussed, the standard deviation over horizon optimal portfolio would be exactly the same portfolio as the minimum monthly standard deviation portfolio if monthly returns were i.i.d. However, if dependencies across time within and between asset classes exist, these criteria measure different things and distinguishing between them may have repercussions for the optimal portfolio. If mean reversion manifested in equity returns (and this effect was reproduced by our modelling) over a given investment horizon, one would expected a higher allocation to equities as it would be relatively more attractive compared with cash and bonds. If, on the other hand, equity momentum is present, the opposite may be the case.

Further, investors may not be primarily averse to return variation around the mean, but instead be averse to underperforming some required level of real return. For example, the outcome of an asset-liability analysis may suggest that an investor requires a certain real return for a financial plan to be viable. Thus, the classical MVO efficient frontier may not be the most relevant paradigm.

Probability of success, average shortfall and downside deviation all penalise underperformance of the required real return, while giving no penalty for variability of returns above the required real return. The probability of success penalty is binary, penalising marginal underperformance and large outperformance exactly the same. The average shortfall penalty is linearly increasing in the underperformance, while the downside deviation is quadratically increasing in underperformance. It is clear that the relative penalty (as a proportion of the total penalty) exacted for large underperformance of required real return in these three measures can be ordered as:

Probability of success penalty < average shortfall < downside deviation < standard deviation

As the penalty is in relation to the required real return rather than the expected real return, there is the distinct possibility that the expected real return of the optimal portfolio may be in excess of the required real return and hence that these portfolios may exhibit materially higher monthly standard deviation than the MVO portfolio.

In contrast to monthly standard deviation and the standard deviation over horizon, the other three parameters above depend on higher-order moments (beyond mean and variance) of the assumed returns distribution. Thus, if we are able to capture any potential non-normality with our modelling choices, it may result in optimal portfolios that are more true to the true distributions of asset returns.

The different risk measures above could also be thought of as representing the different parties involved in financial planning and fund management: (1) the asset management business/individual asset manager; (2) the investor's true interest (say the astute financial planning investor); and (3) the investor's interests.

One could argue that for an asset management business with balanced funds to be successful, it needs to deliver on the real return targets of their funds in the long term. If it does not, it will arguably not survive. The implicit penalty function is perhaps similar in nature to the binary probability of success.

The investor who is advised to enter a fund based on a real return requirement (typically the internal rate of return that balances assets and liabilities) is quite clearly in reality most averse to underperformance of his required real return as it would lead to his or her financial plan not coming to bear. Typically these investors have a very long investment horizon, and their financial advisors coach them to be less averse to month-on-month variability of returns. His or her penalty function will arguably resemble that of the average shortfall and downside deviation portfolio.

However, the typical investor is psychologically (as opposed to financially) arguably very much averse to month-on-month variability. Unsophisticated investors are easily prompted to act on short-term fluctuations experienced in their portfolios. One would expect this type of investor to be most comfortable, for any given level of real return, in the monthly standard portfolio. As a result, many asset managers will also manage their assets in accordance with this same aversion to monthly fluctuations.

It is important to keep in mind the inherent conflicts between these different parties (the asset manager, the financial planning client and the psychologically typical investor), when the resulting optimal portfolios are discussed.

4. DATA AND METHODOLOGY

This chapter is divided into four sections. In the first section (4.1) all the data used in this work is introduced. The second section (4.2) introduces all the concepts needed to define our models, which the third section (4.3) then proceeds to do. The detail of the numerical technique employed to find the optimal portfolios is relegated to appendix 7.1.

4.1. The data

The data employed in this research will be used solely to describe the joint distribution of returns. There are two categories of data:

(1) Monthly asset class returns data; and
(2) Regime-identifying variables.

The monthly asset class returns data will be used to define the unconditional (regime-ignorant) empirical distributions or, when we incorporate regimes, the conditional empirical distributions. The regime-identifying variables will in turn be used to define states (or regimes) as well as state-changing mechanisms. Combined, these two types of data will help us to model the regime-cognisant return distributions of portfolios of assets, comparable to the methods found in typical regime-switching methods.

The returns data include monthly returns for the all the main asset classes found in most SA pension funds: SA equities, SA bonds, SA cash, SA property, global equities, global bonds, global cash and global property. Over the last few years it has become typical for SA inflation-linked bonds to be considered an asset class in its own right. However, it has a far shorter history (2000 to present) than the other asset classes. The benefits of a longer dataset were deemed to outweigh the benefits of including this asset class. As a result it was excluded from consideration in this work.

The regime-identifying variables include the first-order change in GDP and inflation as well as a selection of market variables. The second-order change in some of these variables is also included. This means that, for example, both the *level* of growth and inflation as well as the *direction* of growth and inflation are incorporated in regime-identification. The choice of variables will be discussed in later sections (see 4.1.3 and 4.2.3.1).

4.1.1. Period

The period considered is the 541 months from 1 March 1972 to 31 March 2017, as it was the longest period for which reasonable quality proxies with monthly returns for all asset classes and economic variables could be identified and sourced. All global returns are converted to the South African rand (ZAR) by applying the relevant exchange rate at each point in time.

4.1.2. Monthly asset class returns data

All returns are total returns, reflecting both the price movements as well as income assumed to be reinvested. For SA equities and SA bonds, there is a heavy reliance on the data and research of Firer & McLeod (1999). Refer to their work for a detailed motivation for the choice of indices for those asset classes. Unless stated otherwise, all of the other series were sourced from I-Net Bridge.

Monthly global bond returns could only be found from 1988 onwards. For the period prior to that, the Dimson-Staunton-Marsh (DMS) (2001) global bond calendar year returns were used. Within calendar years prior to 1988, the monthly returns of US 10 year bonds were used, but a constant (within each calendar year) adjustment is made to each month in order to retrieve the calendar year return of DMS. In other words, the average level of monthly returns for that period reflect the global calendar year returns of DMS, while the monthly variability within any one year will reflects that of US 10 year bonds.

Table 2 below shows the data sources for each asset class:

SA equities	
1. Rand Daily Mail Industrial Index	1972 to 1978
2. JSE Actuaries Index from	1978 to 1995
3. South African JSE All Share Index (ALSI)	1995 to 2017
SA bonds	
1. Total return based on a theoretical one bond portfolio, based on the most recently issued 20-year bond's yield and price.	1972 to 1980
2. Composite of JSE Actuaries Bond Indices (0-3, 3-7,7-12,12+ years)	1980 to 1985
3. JSE Actuaries Bond Index	1986 to 1998
4. South African All Bond Total Return Index (ALBI)	1999 to 2017
SA cash	
1. Alexander Forbes Money Market Index	1972 to 2000
2. South African Short-term Fixed Interest (3-month) Benchmark	2000 to 2017
SA property	
1. JSE Property Index	1972 to 1976
2. Market cap weighted between the JE Property Index and the Property Unit Trust Index	1976 to 1990
3. Market cap weighted between the JSE Property Index, Property Unit Trust and Property Loan Stock Indices	1990 to 1999
4. Market cap weighted between SA Property Unit Trust and SA Property Loan Stock Indices	1999 to 2002
5. SA Listed Property Index	2002 to 2017
Global equities	
1. MSCI World Total Return	1972 to 1988
2. MSCI All Countries	1989 to 2017
Global bonds	
1. Dimson, Staunton and Marsh (2001) global bonds calendar year returns with monthly returns within calendar years based on the monthly yield changes of 10 year US bond total return.	1972 to 1988
2. Barclays Global Aggregate Index	1988 to 2017
Global cash	
1. US 1 year interest rate is used as a proxy for the cash rates of all three regions. The three relevant exchange rates are used to convert returns to ZAR The final return is 60% US cash, 30% euro cash and 10% UK cash.	1972 to 2017
Global property	
1. FTSE NAREIT US Real Estate Index	1972 to 1990
2. UBS Global Property Investors Index	1990 to 2017

Table 2: Data sources for asset class returns

4.1.3. Regime identifying data (March 1972 to March 2017)

- SA Real GDP Growth (y-o-y[12])
- SA CPI (y-o-y)
- US GDP Growth (y-o-y)
- US CPI (y-o-y)
- 12-month change in SA Real GDP (y-o-y)
- 12-month change in SA CPI (y-o-y)
- 12-month change in US GDP (y-o-y)
- 12-month change in US CPI (y-o-y)
- ALSI price index (y-o-y)
- S&P 500 price index (y-o-y)
- ALSI real earnings yield
- S&P 500 real earnings yield
- USD/ZAR (y-o-y)

The GDP data are all updated quarterly, three months in arrear, to ensure only information that was actually available was used when modelling returns distributions at any point in time. Similarly, inflation figures are lagged by one month, while market variables are not lagged at all. While this creates a slight misalignment between the date assigned to data on the one hand and the periods that they actually describe, it helps to avoid deriving optimal portfolios with information that was not yet available at the time.

4.1.4. The historical risk and return characteristics of the asset classes

The most important inputs of traditional mean-variance optimisation are the expected returns and the covariance matrix of the assets under investigation. Even though this work is not confined to traditional MVO, these characteristics are similarly important in the current context. It is important that these characteristics are, within our dataset, plausible and broadly representative of what could be expected to come to bear in the long-term future. Table 3 and Table 4 below show the returns[13], standard deviations and correlations that occurred over the dataset, which consists of the 541 months from 1 March 1972 to 31 March 2017.

4.1.4.1. Average return, standard deviations and correlations

	SA equities	SA bonds	SA cash	SA listed property	Global equities	Global bonds	Global cash	Global property	SA CPI
nominal return p.a.	16.96%	11.19%	10.80%	15.69%	14.79%	12.85%	7.22%	14.53%	9.20%
real return p.a.	7.76%	1.99%	1.60%	6.49%	5.58%	3.65%	-1.98%	5.33%	0.00%
annualised monthly std	21.05%	8.15%	1.31%	18.78%	17.19%	14.39%	13.06%	18.57%	2.49%

Table 3: Return and risk of asset classes for the period 1 March 1972 to 31 March 2017

[12] Year-on-year, i.e. the 12-month change.
[13] Natural logarithm returns

	SA equities	SA bonds	SA cash	SA listed property	Global equities	Global bonds	Global cash	Global property	SA CPI
SA equities	1.0	0.3	0.0	0.5	0.4	-0.1	-0.1	0.2	0.1
SA bonds	0.3	1.0	0.1	0.4	-0.1	-0.2	-0.3	0.0	0.0
SA cash	0.0	0.1	1.0	0.0	0.1	0.2	0.1	0.0	0.2
SA listed property	0.5	0.4	0.0	1.0	0.1	-0.2	-0.2	0.1	0.0
Global equities	0.4	-0.1	0.1	0.1	1.0	0.6	0.6	0.7	0.0
Global bonds	-0.1	-0.2	0.2	-0.2	0.6	1.0	0.9	0.6	0.0
Global cash	-0.1	-0.3	0.1	-0.2	0.6	0.9	1.0	0.5	0.0
Global property	0.2	0.0	0.0	0.1	0.7	0.6	0.5	1.0	0.0
SA CPI	0.1	0.0	0.2	0.0	0.0	0.0	0.0	0.0	1.0

Table 4: Correlations between asset classes based on monthly returns from 1 March 1972 to 31 March 2017

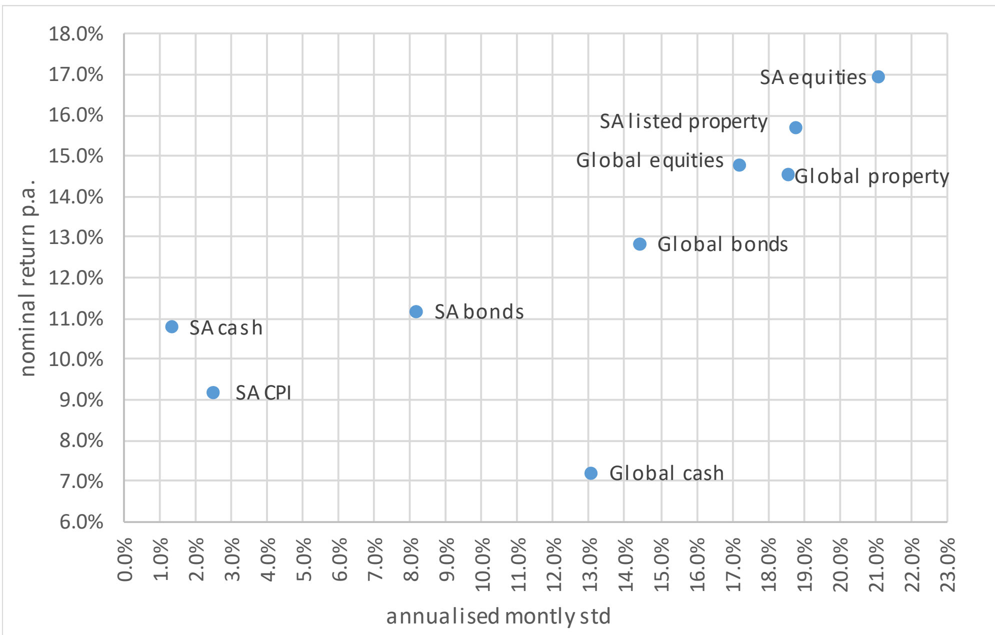

Figure 1: Nominal asset returns versus monthly standard deviations (both in annual terms) for the period 1 March 1972 to 31 March 2017

For the most part the positions of the asset classes in Figure 1 are broadly consistent with theory in terms of their relative risk-return trade-offs. SA cash, SA bonds, SA property and SA equities, in that order, exhibit increasing return and risk, broadly consistent with modern portfolio theory. For the global asset classes a similar progression is noticeable – one surprise is that global property exhibits a higher standard deviation than global equities. Listed property is often perceived to be a hybrid between equities and bonds, but this cursory glance suggests it is far more equity-like than bond-like. It is perhaps not too surprising as the underlying securities are, after all, listed on stock exchanges. The high variability of property returns is likely also a function of the gearing typically employed in property assets.

Finally, it is noted that the South African rand weakened by 6.4% over the period of our analysis.

4.1.4.2. Periodicity-dependent characteristics of asset class returns

In this section some of the features exhibited by asset class returns are discussed. There will be a focus on the most useful phenomena for explaining the optimal portfolios derived for the various criteria, investment horizons and block lengths.

Four possible measures of risk are considered in this section: annualised standard deviation of returns; the annualised standard deviation of real returns; the average shortfall relative to SA CPI; and the downside deviation relative to SA CPI (in the terminology of our definitions of these terms, the required real return threshold will be zero, unless specifically stated otherwise). All of these measures will be calculated employing data of increasing periodicity (i.e. longer units of data in terms of months). All results are again for the period 1972 to 2017, unless specifically stated otherwise.

SA equities versus other asset classes

Figure 2 and Figure 3 below show the annualised monthly standard deviations of returns in ZAR as a function of periodicity for all eight asset classes in nominal and real terms, respectively.

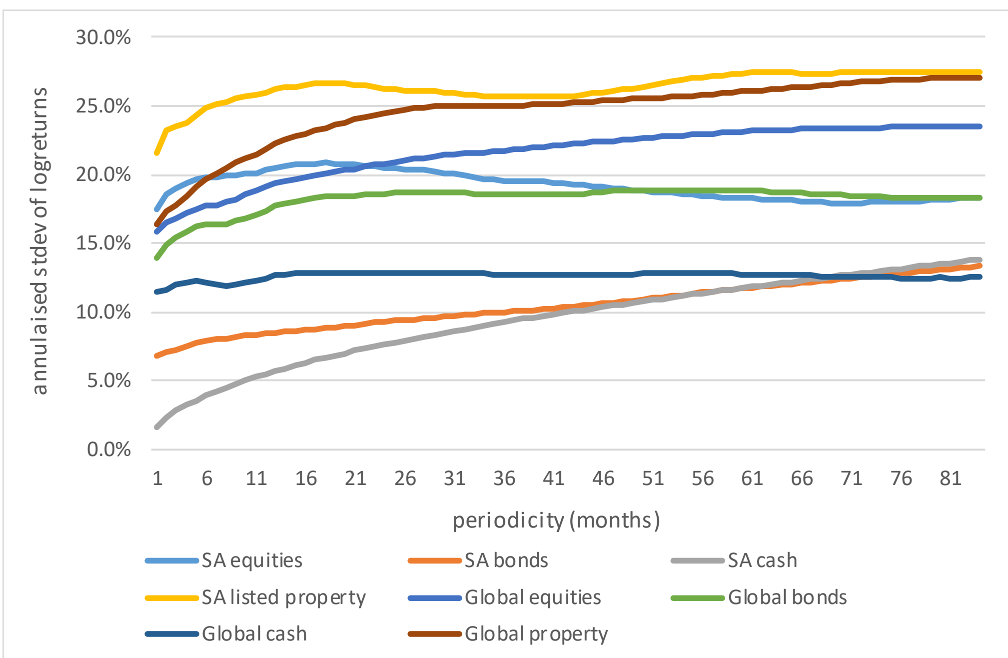

Figure 2: Annualised standard deviations of all asset classes (nominal terms) versus return-periodicity

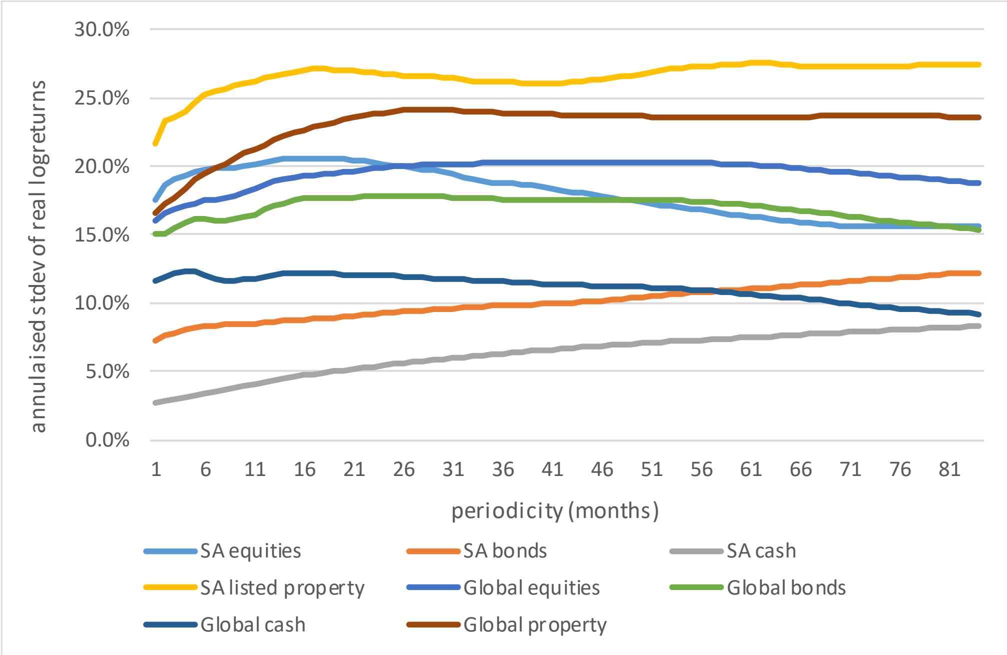

Figure 3: Annualised standard deviations of all asset classes (real terms) versus return-periodicity

The annualised sample standard deviation of nominal returns for a periodicity of length T months is calculated as:

$$\sqrt{\frac{12}{T} * \frac{1}{N_T - 1} \sum_{i=0}^{N_T - 1} (r_{iT,(i+1)T} - \bar{r}_{T,\,N_T})^2} \qquad (45)$$

where $r_{0,T}, r_{T,2T}, \ldots, r_{(N_T-1)T,N_T T}$ are the N_T consecutive (overlapping) observed T-month returns and $\bar{r}_{T,\,N_T}$ is their mean. The annualised sampled standard deviation of real returns is similarly defined, with nominal returns and their mean substituted by the real equivalents.

If nominal returns and real returns were i.i.d., the annualised standard deviations of each would be independent of T, and the above graphs would tend to be straight horizontal lines as the numbers of months of data at our disposal tended to infinity.

Equities have long been purported to exhibit momentum in the short term (approximately 12 months) and mean reversion in the long term (36 to 60 months). If returns were distributed i.i.d. over time, as is often assumed in mean-variance optimisation, one would expect the annualised standard deviation of returns calculated over different periodicities to be the same. It is only the standard deviation of SA equities in Figure 2 and Figure 3 above that exhibits a relatively consistent decline over increasing periodicities.

For clarity, Figure 4 and Figure 5 below show only the standard deviation of SA equities, SA bonds and SA cash, in nominal and real terms, respectively:

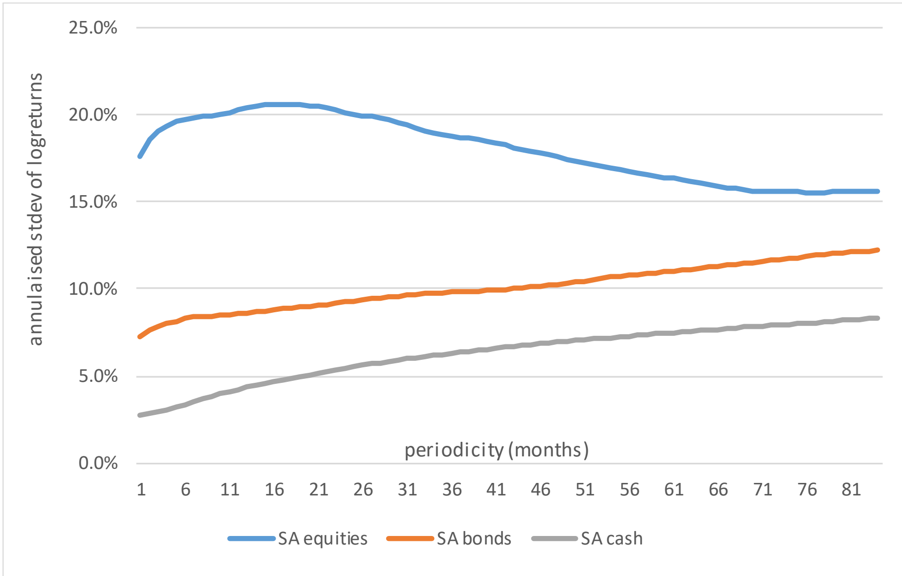

Figure 4: Annualised standard deviation of nominal returns for SA equities, SA bonds, and SA cash versus periodicity

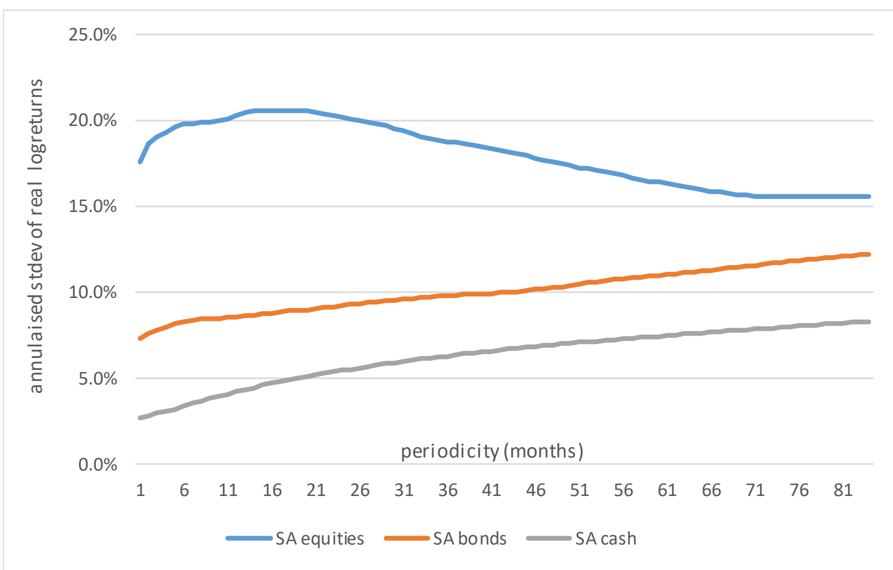

Figure 5: Annualised standard deviation of real returns for SA equities, SA bonds, and SA cash versus periodicity

From Figure 4 and Figure 5 it is evident that the SA equities annualised standard deviation initially increases, peaking at around 20 months, broadly consistent with the notion of short-term momentum. However, it then decreases as the periodicity increases, as one would expect if long-term mean reversion existed. The higher the periodicity, the fewer the mutually exclusive number of data points – it is possible that the increase after around 71 months is due to the small sample, rather than an actual characteristic of SA equities. On the other hand, the standard deviation of SA bonds and SA cash are monotone increasing – this makes sense if one considers the effect of low and high interest rate regimes.

Figure 6 and Figure 7 show the ratios of standard deviations of the various asset class in respectively nominal and real terms. Despite the initial increase in the SA equities annualised standard deviation between months 1 and 20, the standard deviation of SA equities as a ratio of the standard deviation of nominal returns of all the other asset classes is for all intents and purposes monotone decreasing over periodicity, though this effect is easily the most prominent against SA cash and SA bonds as Figure 6 and Figure 7 below show:

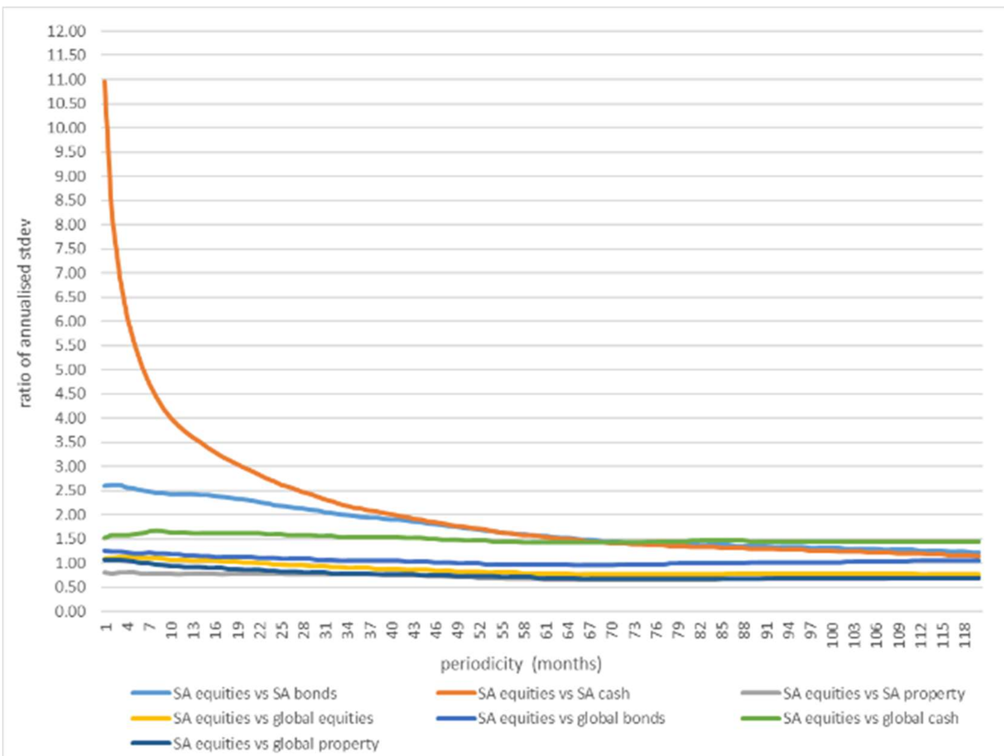

Figure 6: Ratios of nominal standard deviations versus periodicity

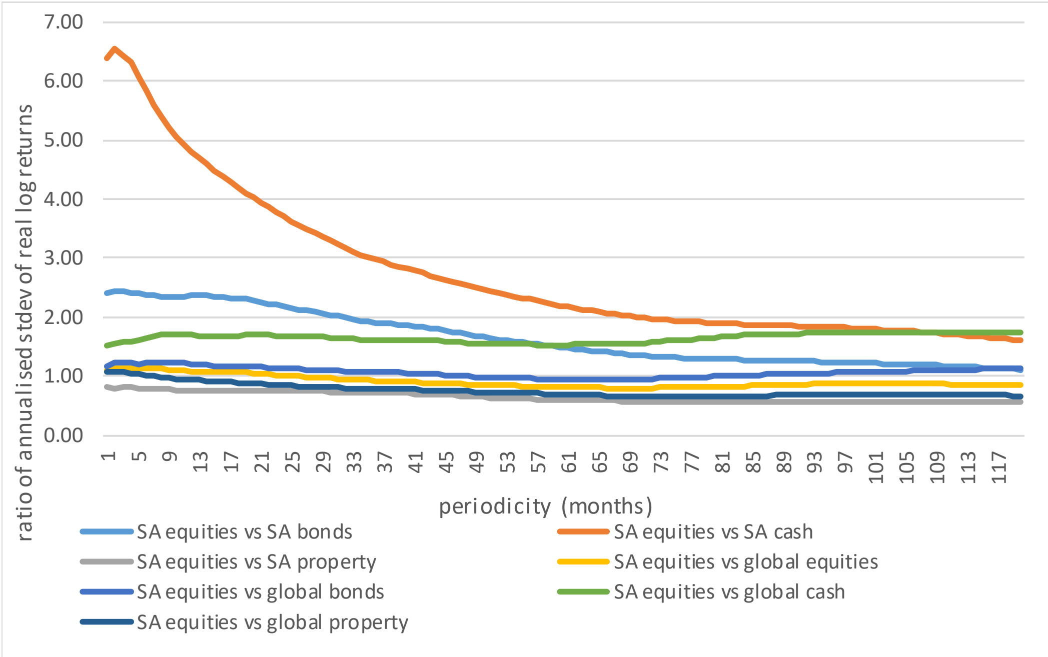

Figure 7: Ratios of real standard deviations versus periodicity

The effects are less consistent but still clearly present for the standard deviation of real returns. If the SA equities versus SA cash ratio is excluded for both the nominal and real ratios of standard deviations (see respectively Figure 8 and Figure 9 below), the other ratios are more easily discernible, and it is clearer that SA equities has a tendency to become relatively less volatile than most other asset classes as periodicity increases:

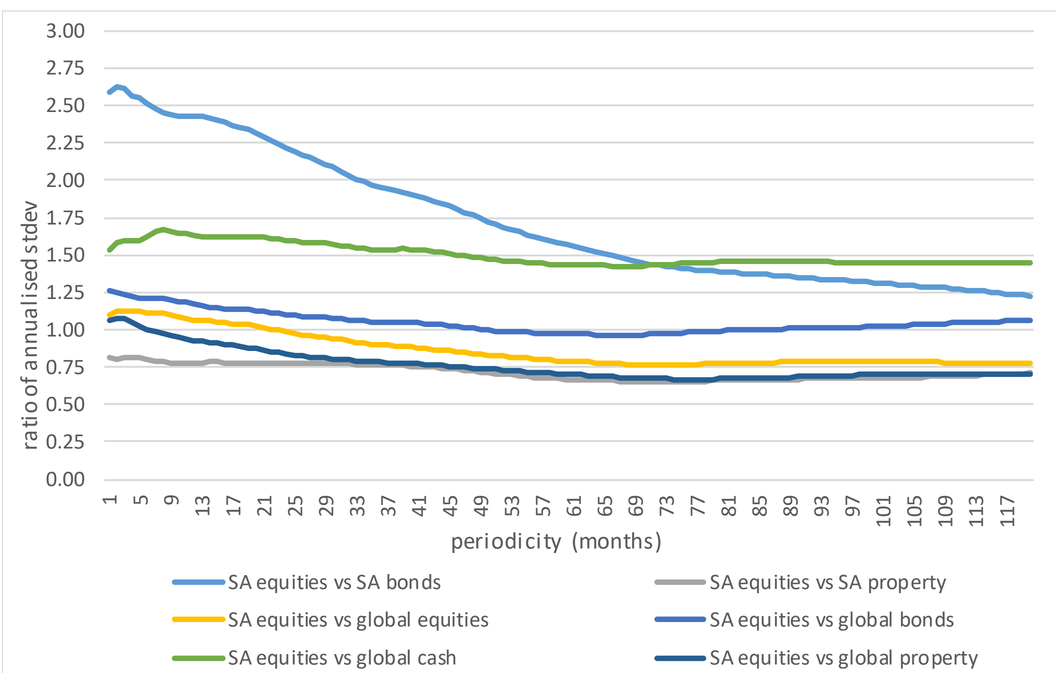

Figure 8: Ratios of nominal standard deviations (excl. local cash) versus periodicity

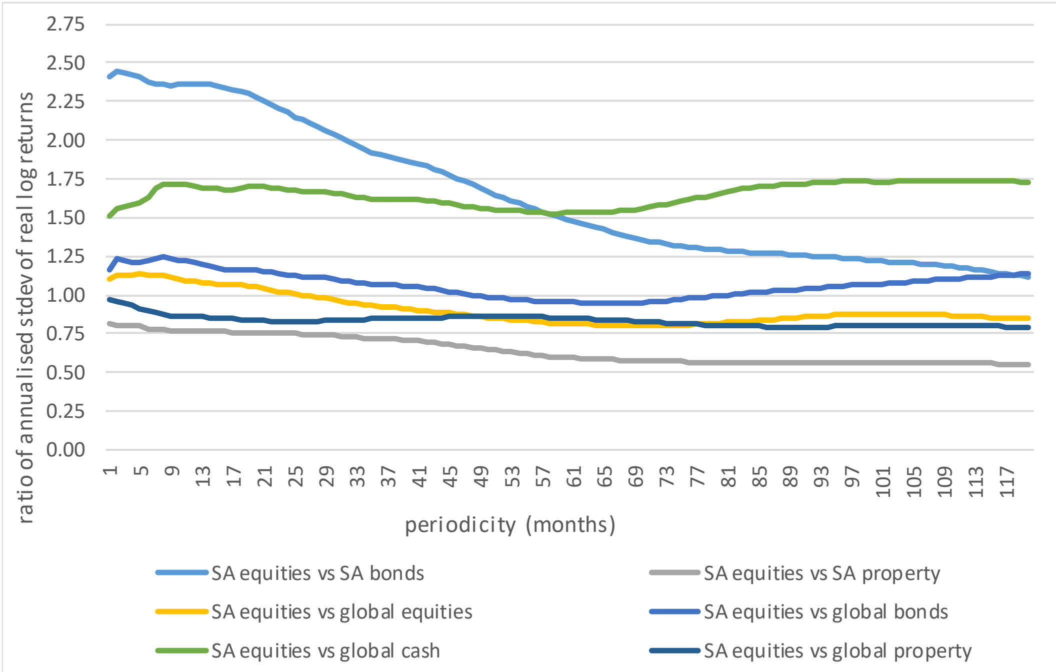

Figure 9: Ratios of real standard deviations (excl. local cash) versus periodicity

SA bonds versus SA cash

As was already evident from the graphs above, the annualised standard deviation of bonds is initially significantly higher than that of SA cash, but for a periodicity of approximately 50 months and longer, the two become virtually indistinguishable. Figure 10 and Figure 11 below show how the ratio of the two standard deviations approaches 1 and 1.5 respectively as periodicity increases:

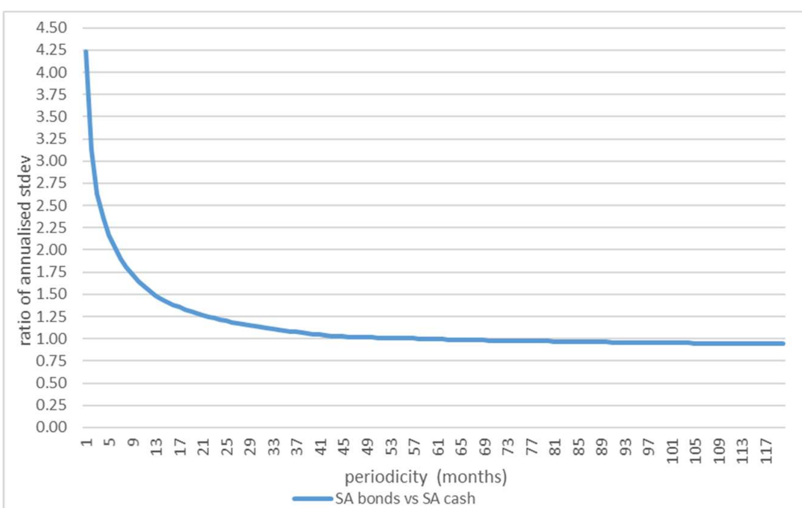

Figure 10: Ratio of standard deviation of SA bonds and cash (nominal) versus periodicity

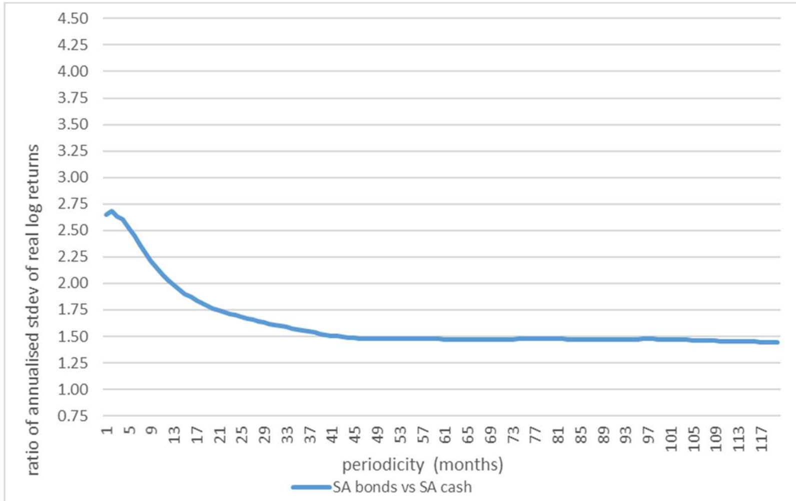

Figure 11: Ratio of standard deviation of SA bonds and cash (real) versus periodicity

Thus, in nominal terms (Figure 10), the standard deviation of the two asset classes converge. In real terms (Figure 11), the ratio reduces *significantly* over periodicity, but cash remains less variable even for the longest periodicities.

Global equities (versus global bonds and global property)

Figure 12 below shows the ratio of the nominal annualised standard deviations of global equities to, respectively, global bonds and global property. Figure 13 shows the same information but in real terms. It is evident that the standard deviation of global equities clearly increases as a ratio of that of global bonds and, less materially, deceases as a ratio of global property. Note that one must not infer that global equities do not exhibit mean reversion: in the respective local currencies of the constituents, they possibly also exhibit long-term mean reversion, however, when converted to ZAR, this effect is perhaps diminished.

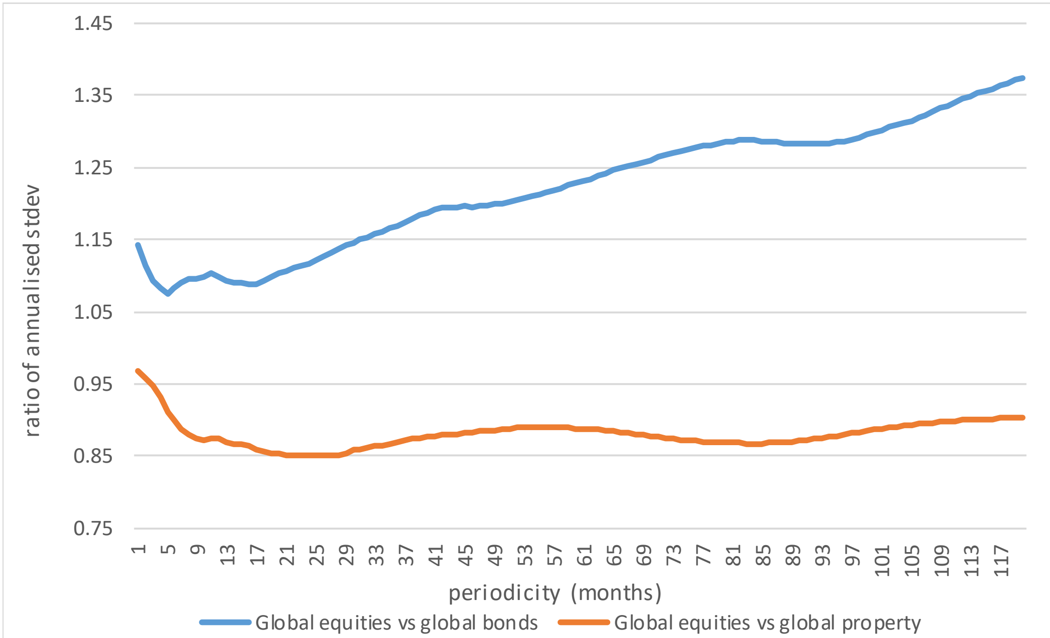

Figure 12: Ratios of standard deviations for global assets (nominal) versus periodicity

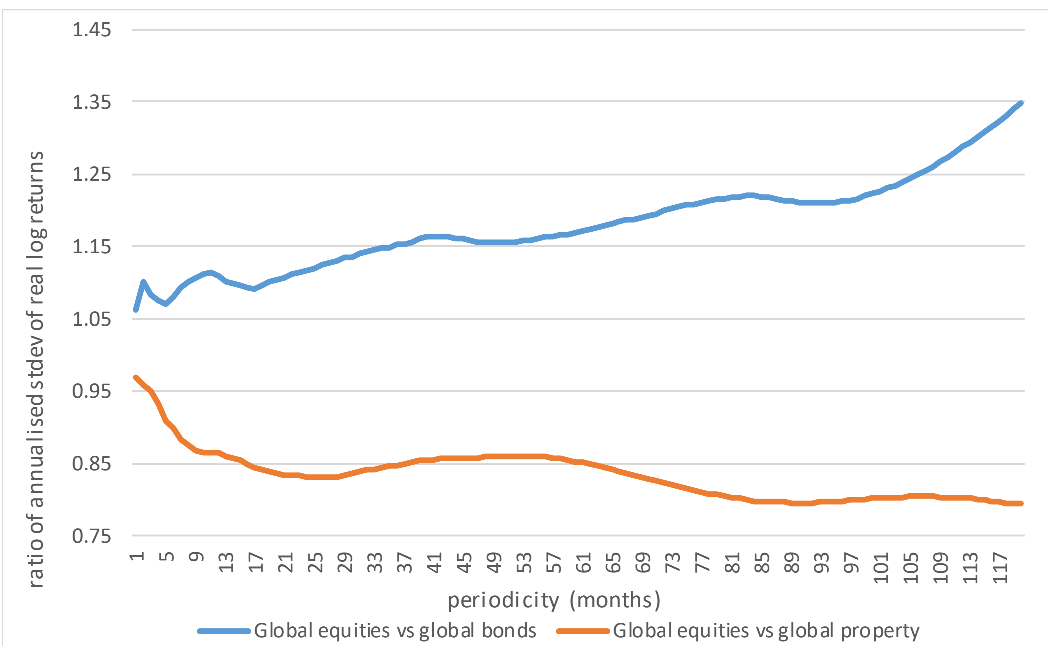

Figure 13: Ratios of standard deviations for global assets (real) versus periodicity

Other measures of risk

The standard deviation measure penalises up and downside variability equally. In reality investors are arguably instead averse to downside variability, as already discussed. If the downside with respect to a return threshold is penalised, the level of real returns of asset classes are also incorporated. Such a measure, when evaluated over various periodicities, could thus indicate the relative attractiveness of each asset class considering both the level and variability of returns, for an increasing periodicity.

Consider again Figure 4 and Figure 5 on page 44, and also Figure 14 and Figure 15 below, the equivalent graphs for the two other risk measures, both evaluated at a real return threshold of 0:

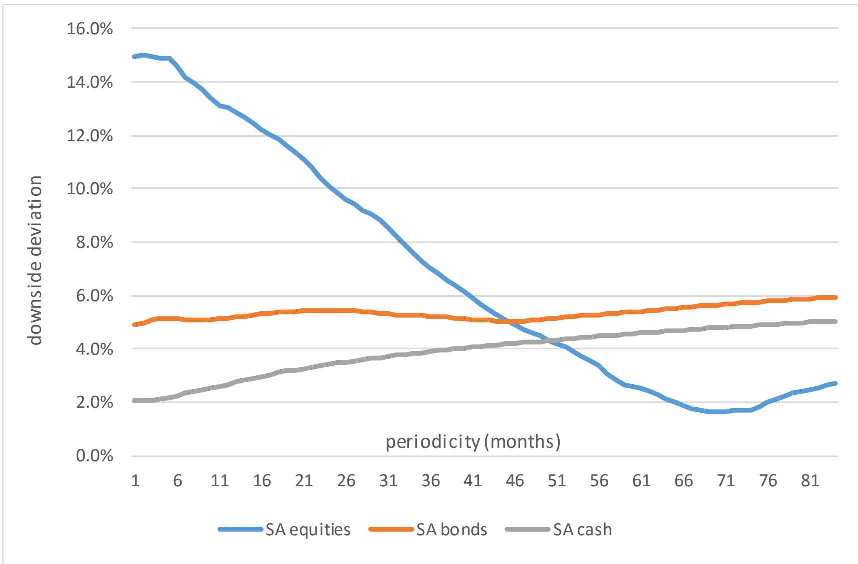

Figure 14: Downside deviation (threshold = CPI) of SA equities, SA bonds and SA cash versus periodicity

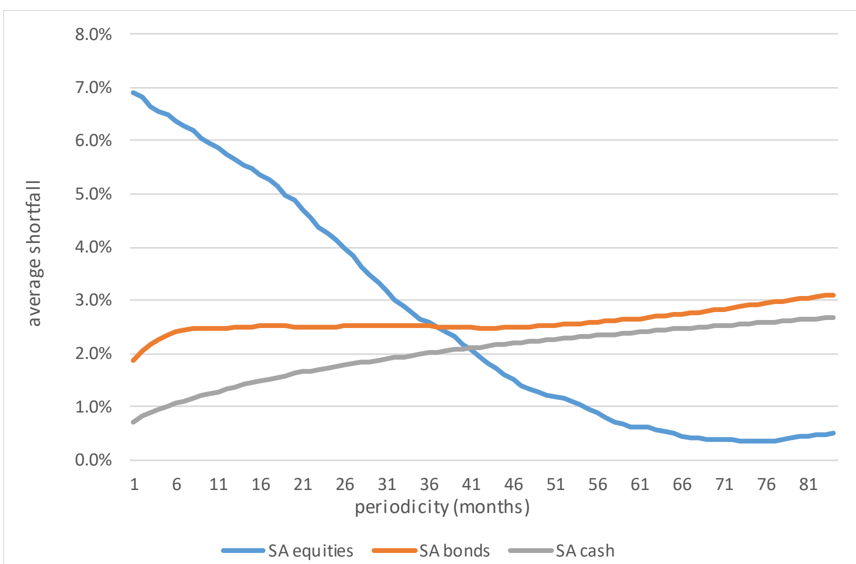

Figure 15: Average shortfall (threshold = CPI) of SA equities, SA bonds and SA cash versus periodicity

In Figure 14 and Figure 15 it is evident that SA equities actually becomes the least risky asset class at around 41 months and 61 months respectively for these two bespoke risk measures: the higher average real return of equities means that it is less and less likely to underperform inflation over longer periods.

In Figure 16 and Figure 17, which show respectively the average shortfall and downside deviation relative to CPI for all the asset classes, it is evident that the only asset classes that exhibit a significant improvement over periodicity are SA equities, global bonds, and global equities, though SA equities have easily the steepest decline:

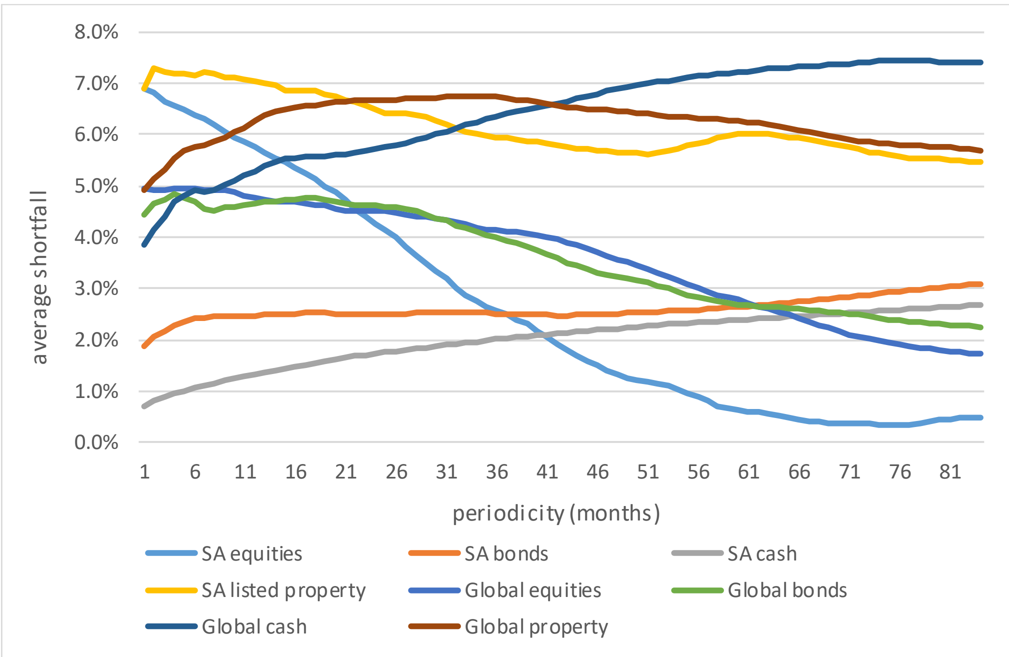

Figure 16: Average shortfall (threshold = CPI) of all asset classes versus periodicity

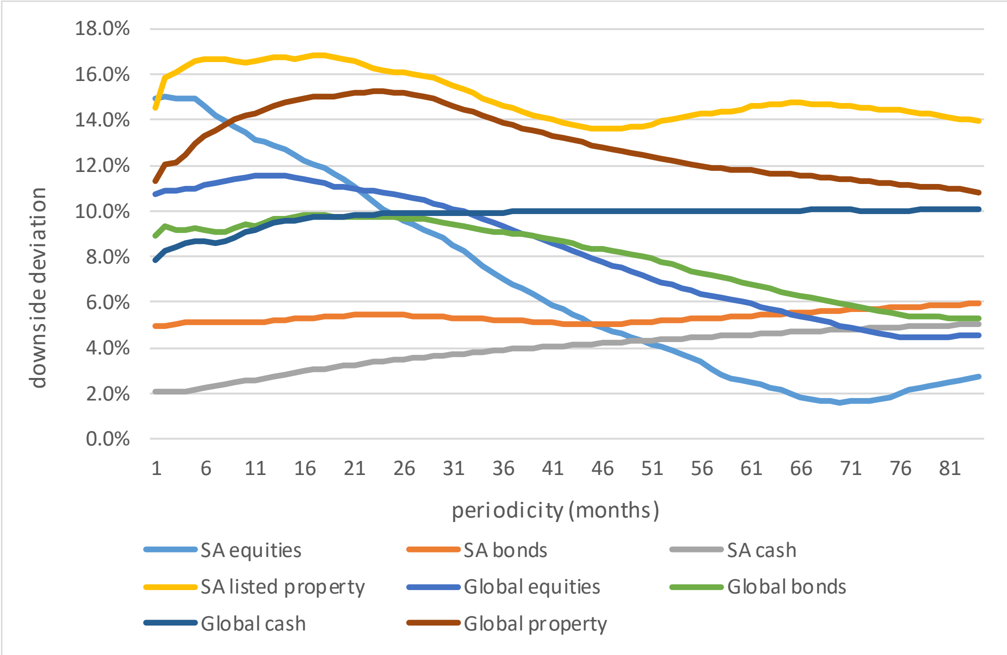

Figure 17: Downside deviation (threshold = CPI) of all asset classes versus periodicity

It is important to realise that the two bespoke risk criteria shown in Figure 16 and Figure 17 are sensitive to the chosen return threshold. While the downside deviation and average shortfall of SA bonds never quite fall below that of SA cash when using CPI as the threshold, the situation changes for higher CPI targets. Table 5 to Table 8 below rank each asset class with respect to the average shortfall, for real return thresholds of 0%, 2%, 4% and 6%, respectively:

period	average shortfall (threshold = CPI+0%) best to worst							
1	SA cash	SA bonds	Global cash	Global bonds	Global property	Global equities	SA property	SA equities
12	SA cash	SA bonds	Global bonds	Global equities	Global cash	SA equities	Global property	SA property
24	SA cash	SA bonds	SA equities	Global equities	Global bonds	Global cash	SA property	Global property
36	SA cash	SA bonds	SA equities	Global bonds	Global equities	SA property	Global cash	Global property
48	SA equities	SA cash	SA bonds	Global bonds	Global equities	SA property	Global property	Global cash
60	SA equities	SA cash	SA bonds	Global bonds	Global equities	SA property	Global property	Global cash
72	SA equities	Global equities	Global bonds	SA cash	SA bonds	SA property	Global property	Global cash
84	SA equities	Global equities	Global bonds	SA cash	SA bonds	SA property	Global property	Global cash
96	SA equities	Global equities	Global bonds	SA cash	SA bonds	SA property	Global property	Global cash
108	SA equities	Global bonds	Global equities	SA cash	SA bonds	Global property	SA property	Global cash
120	SA equities	Global bonds	Global equities	SA cash	SA bonds	SA property	Global property	Global cash

Table 5: Least to most risky asset classes according to average shortfall (threshold = CPI) versus periodicity (high to low)

average shortfall (threshold = CPI+2%)

period	**best to worst**							
1	SA cash	SA bonds	Global cash	Global bonds	Global property	Global equities	SA property	SA equities
12	SA cash	SA bonds	Global bonds	Global equities	Global cash	SA equities	Global property	SA property
24	SA cash	SA bonds	SA equities	Global equities	Global bonds	Global cash	SA property	Global property
36	SA cash	SA bonds	SA equities	Global bonds	Global equities	SA property	Global cash	Global property
48	SA equities	SA cash	SA bonds	Global bonds	Global equities	SA property	Global property	Global cash
60	SA equities	SA cash	SA bonds	Global bonds	Global equities	SA property	Global property	Global cash
72	SA equities	Global equities	Global bonds	SA cash	SA bonds	SA property	Global property	Global cash
84	SA equities	Global equities	Global bonds	SA cash	SA bonds	SA property	Global property	Global cash
96	SA equities	Global equities	Global bonds	SA cash	SA bonds	SA property	Global property	Global cash
108	SA equities	Global bonds	Global equities	SA cash	SA bonds	Global property	SA property	Global cash
120	SA equities	Global bonds	Global equities	SA cash	SA bonds	SA property	Global property	Global cash

Table 6: Least risky to most risky (left to right) asset classes according to average shortfall (threshold = CPI+2%) versus periodicity (high to low).

average shortfall (threshold = CPI+4%)

period	**best to worst**							
1	SA cash	SA bonds	Global cash	Global bonds	Global equities	Global property	SA equities	SA property
12	SA cash	SA bonds	Global equities	Global bonds	SA equities	Global cash	Global property	SA property
24	SA cash	SA bonds	SA equities	Global equities	Global bonds	SA property	Global property	Global cash
36	SA equities	SA cash	SA bonds	Global equities	Global bonds	SA property	Global property	Global cash
48	SA equities	Global bonds	Global equities	SA bonds	SA cash	SA property	Global property	Global cash
60	SA equities	Global bonds	Global equities	SA bonds	SA cash	SA property	Global property	Global cash
72	SA equities	Global equities	Global bonds	SA bonds	SA cash	SA property	Global property	Global cash
84	SA equities	Global equities	Global bonds	SA bonds	SA property	SA cash	Global property	Global cash
96	SA equities	Global equities	Global bonds	SA property	SA bonds	SA cash	Global property	Global cash
108	SA equities	Global bonds	Global equities	SA property	Global property	SA bonds	SA cash	Global cash
120	SA equities	Global bonds	Global equities	SA property	Global property	SA bonds	SA cash	Global cash

Table 7: Least risky to most risky (left to right) asset classes according to average shortfall (threshold = CPI+4%) versus periodicity (high to low)

average shortfall (threshold = CPI+6%)

period	best to worst							
1	SA cash	SA bonds	Global cash	Global bonds	Global equities	Global property	SA equities	SA property
12	SA cash	SA bonds	Global equities	Global bonds	SA equities	Global property	SA property	Global cash
24	SA equities	SA bonds	SA cash	Global equities	Global bonds	SA property	Global property	Global cash
36	SA equities	Global equities	Global bonds	SA bonds	SA cash	SA property	Global property	Global cash
48	SA equities	Global bonds	Global equities	SA bonds	SA cash	SA property	Global property	Global cash
60	SA equities	Global equities	Global bonds	SA bonds	SA property	SA cash	Global property	Global cash
72	SA equities	Global equities	Global bonds	SA property	SA bonds	SA cash	Global property	Global cash
84	SA equities	Global equities	Global bonds	SA property	SA bonds	Global property	SA cash	Global cash
96	SA equities	Global equities	Global bonds	SA property	Global property	SA bonds	SA cash	Global cash
108	SA equities	Global bonds	Global equities	SA property	Global property	SA bonds	SA cash	Global cash
120	SA equities	Global bonds	Global equities	SA property	Global property	SA bonds	SA cash	Global cash

Table 8: Least risky to most risky (left to right) asset classes according to average shortfall (threshold = CPI+6%) versus periodicity (high to low)

Table 5 to Table 8 are interesting in and of themselves, as they can be used to quickly ascertain which asset class is best suited for targeting a given real return over an investment horizon of between 0 and 10 years.

For the moment, the following key points are highlighted:

- SA cash dominates SA bonds for CPI+0% and CPI+2% for all periodicities, but the situation reverses for the higher real returns thresholds when the periodicity increases beyond two or three years. For the two highest real returns thresholds, SA bonds become less risky by this measure than SA cash by 48 months and 36 months, respectively.
- For longer periodicities, global bonds fair surprisingly well at delivering real returns, by this measure.
- Global property and global cash are relatively consistently the worst at delivering real returns of all four levels.
- SA equities dominates SA property for all periodicities.
- SA equities becomes the least risky asset class by 48 months, 48 months, 36 months, and 24 months, respectively, for the different real return targets.

Another way to emphasise some of the points made above is by comparing rolling shorter term performance to rolling long-term performance. Table 18 and Table 19 below show the rolling 12-month and 240-month annualised returns of the three local asset classes:

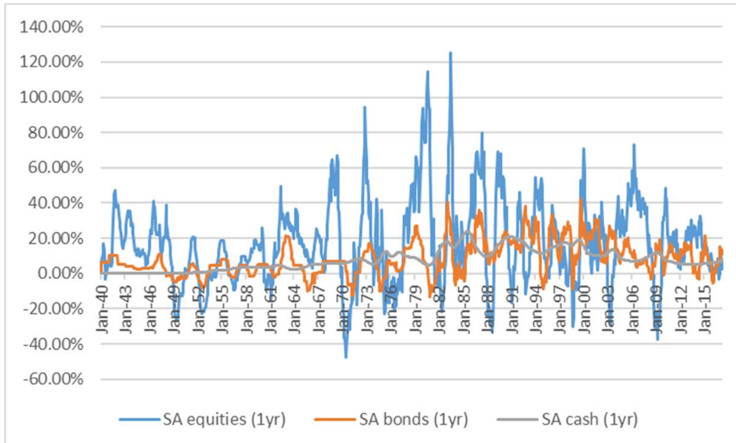

Figure 18: Rolling 12-month annualised nominal returns

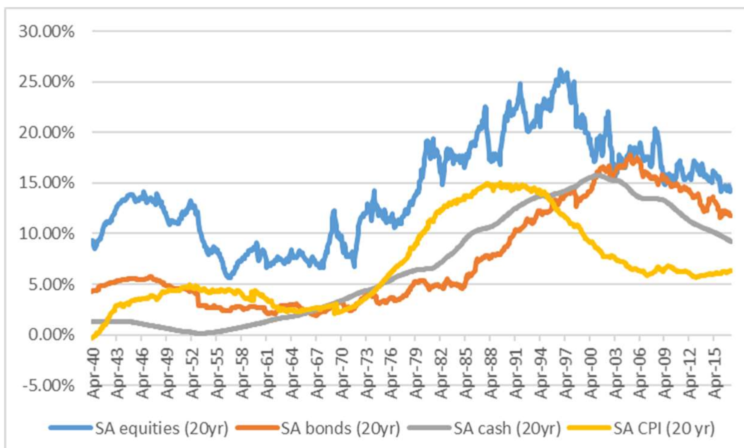

Figure 19: Rolling 240-month annualised nominal returns

Figure 18 shows that the higher short-term variance of equity means makes it less attractive in the short term. However, in the very long term (Figure 19), it is still more variable than the other asset classes, yet equity has nonetheless consistently outperformed bonds and cash (partly due to its higher overall return, and partly due to a declining variability ratio to cash and bonds). SA bonds are significantly more variable in the shorter term than cash, but when measured over the scale of 20-year returns, their variabilities appear very similar from a practical standpoint. Figure 19 also clearly displays the ability of SA equities to consistently deliver real returns in the longer term – the shape of all three asset classes' rolling returns resemble the shape of the rolling SA CPI, suggesting that there is a relationship, even if it is not apparent in short-term data. However, only SA equities consistently delivers above CPI returns. Consider, for example, the correlations between these asset classes since 1925 and inflation for different periodicities shown in Table 9:

months	SA equity	SA bonds	SA cash
1	0.12	0.05	0.32
6	0.12	0.07	0.56
12	0.13	0.20	0.62
24	0.19	0.31	0.67
36	0.28	0.37	0.70
48	0.39	0.43	0.74
60	0.48	0.46	0.76
72	0.56	0.48	0.79
84	0.58	0.49	0.80

Table 9: Correlations (between asset class returns and CPI) versus periodicity

While some of this correlation might be due to overfitting to a small sample (there are only 13 mutually exclusive periods of 84 months in the dataset), it seems clear from Table 9 that these relationships depend on the periodicity. It is not immediately obvious what the implications are for the optimal portfolio, but we will attempt to capture or partially take into account these effects in our optimal allocations, discussed in Chapter 5. However, it is clear that MVO based on shorter periodicities (which is typically the case in practice), fails to account for these longer term relationships.

Summary of most important findings from this section:

- SA equities exhibit material short-term momentum (peaking at around 17 months) and mean reversion over periodicities of between 17 months and 70 months, improving their attractiveness by all measures against all other asset classes as the periodicity is increased. However, *relative* to other asset class, they are virtually monotone declining in risk.
- SA bonds become relatively less variable versus SA cash in terms of annualised standard deviation over increasing periodicity (in both real and nominal terms).
- Global bonds become relatively less variable compared with global equities in terms of annualised standard deviation over increasing periodicity (in both real and nominal terms).
- Global equities, in turn, become less variable versus global property over increasing periodicity (in both real and nominal terms).
- In terms of downside deviation and absolute shortfall with real return thresholds, SA equities, global bonds, and global equities are, for longer investment horizons, the least risky asset classes, even though they are some of the most volatile asset classes in the short term.
- SA equities, SA bonds and SA cash have a clear, positive relationship with SA inflation, and this relationship manifests increasingly over increasing periodicities.

4.2. Concepts employed to define the model: The empirical distribution, block bootstrap and regime-classification

In this section the conceptual building blocks of our models are introduced and discussed: the empirical distribution, block bootstrap methods and the regime-classification methodology.

As already discussed in the literature review, some approaches to modelling returns with regimes can be thought of as dividing non-normal returns into groupings within which normality is approximately retrieved. To what extent the returns distributions within these groupings are in fact normally distributed, or satisfactorily close to normally distributed, is a question that is not typically explicitly addressed by these studies. It is thus not clear how successful the various methods are in this endeavour and, if unsuccessful, what the implications are for model validity and long run performance of these models.

In contrast, in this thesis we do not employ the normal distribution to model the distribution of returns. As a result there is no requirement that normality is retrieved within regimes. Instead, two broad methods are employed in combination to model the shape of the returns distributions over different periods and periodicities:

(1) Simulation from the empirical distribution of returns over some periodicity (which will be defined and referred to as the block length in later sections); and
(2) The placing together of blocks of data simulated in (1) with reference to a regime state.

The returns are not required to be normally distributed within blocks. In fact, each and every asset class and even portfolio will have its own, unique resultant empirical distribution that is highly responsive to its underlying data and, when we incorporate regimes, is conditional on the regime state.

In effect our final models, which are formally defined in Section 4.3, can be thought of as generalisations of the empirical distribution in two distinct senses:

(1) The probability of resampling will deviate from equal weighting to obtain a distribution conditional on the current regime; and
(2) The block length (or the periodicity of the data unit) will be allowed to be longer than one month.

Both of these generalisations (as well as other features of our final model) were inspired by block bootstrap methods. As a result this section will, after the introduction of the empirical distribution, give a brief introduction to block bootstrap methods. Lastly a description and discussion follow on the methodology employed to arrive at returns distributions that account for regimes.

4.2.1. The empirical distribution

In essence the use of the empirical distribution assumes that the observed data are i.i.d. It also, in effect, assumes that each observed return is equally likely to occur again in future. This distribution is extremely responsive to the data at hand, broadly applicable, and hence versatile in its ability to model any distribution under the assumption that the observations are independent and identically (but not necessarily normally) distributed.

4.2.1.1. Definition of the empirical distribution

Let $X_1, X_2, \dots, X_n$ be independent and identically distributed random variables. The cumulative distribution function can be defined as[14]:

$$F_n(x) \equiv \frac{1}{n}\sum_{i=1}^{n} I(x_i < x) \tag{46}$$

where *I* is the indicator function. The empirical density function can be written as

$$f_n(x) = \frac{1}{n}\sum_{i=1}^{n} \delta(x_i < x) \tag{47}$$

where δ is the Dirac delta. The Dirac delta is defined by

$$\int_{-\infty}^{x_i} \delta(x_i < x) = \begin{cases} 0 & if\ x_i < x \\ 1 & if\ x_i \geq x \end{cases} \tag{48}$$

This empirical density function can easily be shown to be consistent with the empirical distribution function:

$$\begin{aligned} \int_{-\infty}^{x} f_n(y)\,dy &= \int_{-\infty}^{x} \frac{1}{n}\sum_{i=1}^{n} \delta(x_i < y)\,dy \\ &= \frac{1}{n}\sum_{i=1}^{n} \int_{-\infty}^{x} \delta(x_i < y)\,dy \\ &= \frac{1}{n}\sum_{i=1}^{n} \int_{\mathcal{R}} I(y \leq x)\delta(x_i < y)\,dy \\ &= \frac{1}{n}\sum_{i=1}^{n} I(x_i < x) \end{aligned} \tag{49}$$

[14] Definitions for the empirical cumulative distribution and density functions from Meucci (2008:28).

$$= F_n(x)$$

4.2.1.2. Properties of the empirical distribution

Some of the attractive properties of the empirical distribution (under the assumption that all data points are i.i.d.) are (Castro, no date):

- **Property 1:** The empirical distribution uniformly converges to the true distribution in sample size:

$$\|F_n(x) - F(x)\|_\infty \xrightarrow{a.s.} 0$$

- **Property 2**: It yields an unbiased estimator of the mean (i.e. is "centred on the data").

- **Property 3**: It asymptotically yields an unbiased estimator of the variance.

Proofs

Property 1: Converges to the true distribution

It follows via the strong law of large numbers that $F_n(x) \xrightarrow[n\to\infty]{a.s.} F(x)$ for every x.[15] *Uniform* convergence follows via the Glivenko-Cantelli theorem. See Van der Vaart (2000:266) for the full proof.

Property 2: Centred on the data

The mean of the distribution is defined as $\mu = \int x dF(x)$. An obvious estimator of the mean is simply $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$. It is trivial to show that the expected value of this estimator is the mean of the distribution, and thus that it is an unbiased estimator of the mean of the distribution:

$$E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = E(X_i) = \mu \quad (50)$$

Property 3:

The variance is defined as:

[15] a.s. is short for "almost surely", i.e. $F_n(x)$ converges to $F(x)$ with probability 1 as $n \to \infty$ for every x.

$$\sigma^2 = \int \left(x - E(x)\right)^2 dF(x) = \int x^2 dF(x) - (\int x dF(x))^2 \quad (51)$$

A natural estimator of the variance of the distribution is

$$\hat{\sigma}^2 = \int x^2 dF_n(x) - (\int x dF_n(x))^2 = \frac{1}{n}\sum_1^n x_i{}^2 - \bar{x}^2 = \frac{1}{n}\sum_1^n (x_i - \bar{x})^2 \quad (52)$$

The last expression is the sample variance, which is well known to be an asymptotically unbiased and consistent estimator of variance.

4.2.1.3. Plug-in principle

The proofs of properties 2 and 3 also demonstrate how the empirical distribution is employed to estimate the distributions of quantities based on the underlying variable via what is called the *plug-in principle*. For the mean and variance, the distributions can be calculated analytically. For more complicated statistics, such as the various risk criteria employed in this current research, simulations are required in the form of repeated withdrawals with replacement from the empirical distribution. If a risk measure can be calculated on a set of returns, it can be simulated in this manner.

4.2.1.4. Discussion of the empirical distribution in the current context

All of the properties of the empirical distribution above are clearly attractive in almost any context. However, the fact that the distribution is centred on the mean of the data is especially important in the context of asset selection. The average return of a portfolio is the most important characteristic of its distribution, and the fact that the empirical distribution retains the average return is important in arriving at practically useful and interpretable results.

The implied versatility of the empirical distribution is a further essential quality in our context. The empirical distribution does not impose any specific shape on the distribution function of the underlying portfolio or on the constituent asset classes. Instead, the returns distribution model will be responsive to the historical distribution of returns of the portfolio at hand. Thus it will be responsive to the portfolio's asset class make-up. In addition, we model each regime-conditional distribution with its own unique empirical distribution, as will become clearer when we state the full distribution models.

Further, since a normal distribution is not imposed on returns, we do not have to constrain ourselves to only mean-variance optimisation. Mean-variance optimisation implicitly assumes normality (or at least that investors are indifferent to non-normality), thereby reducing the concept of risk to variance only, and reducing asset selection to selecting a point on the mean-variance efficient frontier. However, if the empirical distribution is employed, skewness, kurtosis, dependencies across time, and other stylised facts can be incorporated into the analysis.

Although computationally potentially expensive and challenging from a programming perspective, the empirical distribution enables meaningful optimisation with respect to almost any conceivable statistic, and in doing so can cater for alternative risk measures and investor utility function even beyond those employed in this work.

4.2.1.5. Kernel density estimation

There are alternatives to the empirical distribution in this current setting (and in bootstrap methods in general). The most prominent of these methods is perhaps kernel density function estimation. This method can overcome some of the oversensitivity to available data that the empirical distribution suffers from in estimating the underlying distribution, by smoothing out some of the potential variance in the estimates, potentially at the cost of inducing some bias in the extreme edges of the distribution.

Kernel density function estimation essentially uses a more complicated kernel (e.g. the normal density) instead of the Dirac deltas (as defined in (48) of Section 4.2.1.1 above) to spread out the information for any one data point to the surrounding range, arriving at a far smoother density function, while still resulting in a pdf that is responsive to the shape of the distribution of observed data.

In the most general terms,

$$f_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_h\,(x - x_i) \tag{53}$$

where h is the bandwidth parameter that determines to what extent the information is spread out or concentrated at each sample point, and K_h is a function that depends on h.

For example, a monthly return of 2% will not increase the density only at the point 2% – it will increase the density estimate everywhere but with its mode at the point 2%, with gradually less and less weight given to points further away from 2%.

For example, a popular choice is the normal density kernel:

$$f_n(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} e^{-\frac{1}{2\sigma^2}(x-x_i)^2} \tag{54}$$

If the normal density is used, the mean will be the value of the sample and the bandwidth will be the standard deviation parameter (not directly related to and not to be confused with the variance of returns), which will spread out the density of a single sample point to [-∞, +∞]. The weight at each point x_i is clearly spread out to all points on [-∞, +∞], with most weight given to the point x_i, the mean of the normal density.

In relation to our setting, the use of a kernel density would still mean that each portfolio (i.e. asset allocation) would have its own unique distribution that is responsive to the data, as opposed to say the shape of the normal distribution density being imposed on the distribution of returns of all portfolios.

However, the use of kernels introduces further assumptions and complexity, typically requiring a choice to be made in terms of the kernel to use and the "bandwidth" parameter. The bandwidth parameter determines to what degree the information in any one sample value is spread out across the surrounding range. The resulting distribution will no longer be unbiased, but it will have greatly reduced variance – i.e. bias is traded off for variance. The higher the degree of smoothing, the higher the resultant bias of estimates and the lower the variance are.

The bandwidth parameter performs a similar function as our block length parameter, which we will discuss in more detail in the following sections.

Mostly for the sake of simplicity, this work contains its focus to the empirical distribution. However, it is important to note that more sophisticated kernels can be used in this setting and may present an opportunity for significant enhancement in further research. The optimisation techniques that follow could equally well be applied if we chose a more complicated kernel.

4.2.2. Bootstrap

In this section block bootstrap methods are introduced. These methods contain useful adaptations of the bootstrap in the context of weakly dependent data, and will help us to generalise the concept of the empirical distributions in order to potentially further improve our estimates of the returns distributions.

The block bootstrap is a resampling technique for weakly dependent, stationary data. Instead of resampling individual observations from a sample as per the standard bootstrap, blocks of data are resampled in an attempt to partially preserve any dependency structure across successive observations in the data generating process. Variations of the method, their relative performance, and the selection of the optimal block length are described and discussed.

Technically, we will not perform the procedure referred to as "bootstrap" in our modelling of returns distributions, as this is not an estimation of the standard errors of estimates of population parameters. We are more interested in its ability to circumvent the need for complicated analytical methods in the modelling of dependence across time. As our application is different from the usual bootstrap applications, a technical overview of bootstrap and block bootstrap methods is omitted, as this would digress from scope of this current research. For a very detailed technical treatment of a variety of block bootstrap methods, see Lahiri (2003).

Our final model does, however, employ the empirical distribution, as is typically the case in bootstrap. As such we will be employing many of the same underlying concepts and mechanisms that are used in bootstrap and block bootstrap methods, such as blocks, block length, circularity and overlapping data, and will briefly introduce these concepts as they arise in the block bootstrap setting.

4.2.2.1. Standard bootstrap

Bootstrapping (Efron, 1979) is the sampling from any distribution that approximates the true distribution to assess the distribution of an estimator of a parameter of the population.

It is simple to calculate point estimates of the mean and variance from a sample, but the distributions of these point estimates relative to their true values are not known. Without further sampling, how can the accuracy or distributions of these point estimates and their relationship to the true values be inferred?

The bootstrap methodology typically goes about addressing these questions in the following manner: treat for a moment the original sample (say of size N) as a population (importantly, one that approximates the true, original population), and draw N times with replacement from this "sample population" until a new "pseudosample" of the same size as the original dataset is created. Repeat this last step as many times as is needed until as many pseudosamples (say B) as required to obtain a sufficiently refined distribution of the pseudosamples are created. By examining the relationship between the original sample, which is assumed to approximate the underlying population, and the pseudosample, a relationship between the true population and the original sample can be inferred.

The bootstrap procedure's popularity and usefulness in estimating the distribution of point estimates is mostly due to the fact that it can be used in a large variety of situations and contexts: the technique requires very few assumptions regarding the distribution of the population and avoids the need for analytical expressions for the distributions of parameters under investigation (which can be difficult to assess in practice).

Typically the empirical distribution is used to approximate the true distribution, as is the case in our description above.[16]

4.2.2.2. i.i.d. assumption and the need for block bootstrap

While the bootstrap procedure requires remarkably few assumptions, there is one quite fundamental assumption implied by the resampling procedure: that the sample points are identically and independently distributed. It is obvious that by drawing random data points from the original dataset and grouping them randomly together, any possible dependency structure between these points are destroyed. The standard bootstrap is

[16] Other candidates are, for example, smoothed estimates of the true distribution based and parametric methods (which assumes some knowledge of the population distribution), or more complicated kernels.

therefore not appropriate in the context of a data generating process that exhibits any dependencies across proximate observations.

One possible method to deal with dependent data is to fit a parametric model on the data and reduce the problem to the i.i.d. case. The most popular choices are possibly autoregressive (AR) models or moving average (MA) models, or a combination of the two (ARMA). However, such parametric methods introduce the complexities of model selection and parameter fitting, and it is difficult to know to what extent the analysis is influenced by the underlying assumptions made.

The first block bootstrap methods were proposed by Hall (1985) in the context of spatial data. Carlstein (1986) and Kunsch (1989) were the first to advocate the use of block bootstrap methods for time series, where observations close to each other in time often have complex dependency structures that are not very well understood.

4.2.2.3. Block bootstrap methods

A block bootstrap procedure resamples blocks of neighbouring data points instead of individual data points, with the aim of preserving the dependency structure within these blocks. The blocks of data (instead of single observations) are now assumed to be i.i.d. The resampled blocks are then placed next to each other and spliced in order to form a pseudosample.

While the dependency is preserved within blocks, clearly any dependency existing over periods longer than the block length is not. Blocks are still randomly placed next to each other in the block bootstrap, thereby causing a discontinuity at the splice[17]. The longer the block length and the weaker the dependency over longer ranges, the more dependency will be preserved by the blocks. However, the longer the original dataset, the longer is the optimal block length, so that, asymptotically, the dependence structure is fully preserved, in theory.

However, for a finite sample, the longer the block length, the fewer blocks there are and the heavier the reliance on the outcome of the original sample. The block length is sometimes referred to as a smoothing parameter, as increasing it generally leads to an increase in the bias and a decrease in the variance of estimates. The optimal block length is sometimes defined as the block length that will result in the lowest mean square error of an estimate of a parameter.

The block length is one of the very few decisions a practitioner of the block bootstrap needs to make. It should be large enough to sufficiently preserve the dependence structure of observations, but short enough for there to be a sufficiently large number of blocks to construct a relatively rich empirical distribution that satisfactorily approximates the true distribution. Unfortunately there are no general methods available to determine the optimal block length.

[17] As will become clear later, the matching block bootstrap attempts to address this issue.

4.2.2.4. Assumptions of block bootstrap

Where the main assumption underlying the standard bootstrap procedure is that the individual sample points are i.i.d., the main assumption of the block bootstrap is, as already mentioned, that the blocks are i.i.d. For blocks to be identical, it is obvious that the underlying data generating process (DGP) must necessarily be stationary across observations:

$$F(x_{t_1+\tau}, x_{t_2+\tau}, \dots, x_{t_k+\tau}) = F(x_{t_1}, x_{t_2}, \dots, x_{t_k}) \tag{55}$$

for any value k and τ.

On the other hand, for blocks to be considered independent it is necessary to assume that the dependence structure between successive observations is sufficiently weak for any dependence to virtually negligible over long intervals. For example, an assumption that is often made when the distribution of the mean of a DGP is under investigation is that “strong mixing” holds (see Lahiri, 2003:46 for a complete definition). According to Lahiri, the most important implications for strong mixing are:

$$\sum_{i=1}^{\infty} |cov(X_1, X_{1+i})| < \infty \tag{56}$$

and hence

$$\sigma_\infty^2 = \lim_{N\to\infty} var(N^{\frac{1}{2}}\bar{X}_N) = var(X_1) + 2\sum_{i=1}^{\infty} cov(X_1, X_{1+i}) < \infty \tag{57}$$

$$\sigma_N^2 \to \sigma_\infty^2 < \infty \text{ in probability} \tag{58}$$

where

$$\sigma_N^2 \equiv var(N^{\frac{1}{2}}\bar{X}_N) = var(X_1) + 2\sum_{i=1}^{N}(1-\tfrac{i}{N})cov(X_1, X_{1+i}). \tag{59}$$

4.2.2.5. The different block bootstrap schemes in the literature

In the descriptions below, the following abbreviations are used: Block length (L) and the number of datapoints (N).

The simplest block bootstrap scheme is the non-overlapped block bootstrap (NBB) first discussed by Hall (1985). The NBB groups data into N/L mutually exclusive blocks of the same length (with the possible exception of the final block, which is the remainder of observations).

It is clear that the entries in the resampled data will not be identically distributed (given the sample): The first observation could only take on the values $(X_1, X_{1+L}, X_{1+2L}, ...)$, whereas the second observation could only take on the values $(X_2, X_{2+L}, X_{2+2L}, ...)$. Given the sample, these two distributions are clearly distinct. This is not ideal, as ideally the pseudosamples should retain the characteristics of the original sample (in this case, stationarity).

The **moving block bootstrap** (MBB) (Hall, 1985) defines $N - L + 1$ overlapping blocks on the sample. Each block overlaps the previous block at all but one of its observations.

This scheme will to some extent, but not fully, remedy the problem of differently distributed observations in a pseudosample pointed out in the NBB. For example, the first observation in the pseudosample can now take on any one of the original observations in our actual sample, except the last $L - 1$. The second observation in a pseudosample can take on any value in our original sample apart from the first and the last $L - 2$ observations, and so on. This is a clear improvement on the NBB.

However, this improvement comes at the cost of a new problem: the first $L - 1$ and last $L - 1$ observations in the original sample appear in less than L blocks, whereas all the observations in the middle appear in exactly L blocks. This means that our pseudosamples will not be "centred on the data" anymore. For example, the expected value of the mean of all our pseudosamples will be different to the mean of the original sample, which is an obvious misrepresentation of our sample and thus a disadvantage.

The **circular block bootstrap** (CBB), introduced by Politis & Romano (1991), elegantly addresses this bias introduced by the MBB by wrapping the original dataset in a circular fashion: the last observation is followed by the first, and the original sequence continues. This allows exactly N blocks to be defined and for each observation in the original sample to appear in exactly L blocks. This means that the resampled pseudodatasets will once again be centred on the original sample.

This benefit comes at the loss of discontinuities in the last few blocks, where observations that are not truly proximate are forced together. However, the number of blocks without this discontinuity will generally vastly outnumber the affected blocks and have little effect on the overall analysis. This is deemed by its creators to be a small price to pay for the benefit of centring the data correctly and rendering the block bootstrap method once again to be more generally applicable.

All of the block bootstrap methods above have one important drawback: the pseudosamples will not be stationary. All of the methods considered above have a fixed block length, and one implication of this is non-stationarity of the resampled pseudosamples. The **stationary block bootstrap** (SBB), first described by Politis & Romano (1994), overcomes this problem by employing a geometrically distributed block length. Instead of randomly selecting blocks of fixed block lengths, we need to generate two random variables to select a grouping of observations:

(i) A number between 1 and N (i.e. discrete uniform) that determines the starting point of the block; and
(ii) A geometrically distributed randomly variable that determines the length of the block.

This is repeated (independently) until there are N observations, to give us one pseudosample. In this scheme we need to refer to the expected block length, which now becomes the all-important parameter to decide. However, since the block length is varied, the stationary block bootstrap has the very important advantage of being less dependent on the expected block length than the other procedures are on their (deterministic and constant) block lengths. Politis & Romano (1994) show that SSB estimates can be viewed as approximately a weighted average of the estimates emanating from varying a block length in a fixed block length scheme.

The last block bootstrap method discussed is the **matching block bootstrap**. This block bootstrap, first described by Carlstein, Hall, Hesterberg & Kunsch (1998), resamples consecutive blocks not with equal probability but instead probabilistically matches blocks at their ends. The idea is to retain, to an extent, some of the dependency structures that exist in the DGP by removing some of the discontinuities created by splicing blocks randomly together. The probability of drawing a block at any point in time will depend on how similar its first entry is to the last entry of the last block drawn – in other words, it is a first-order Markov chain where the state is changed to the end of the last drawn block, and the transition probability depends on how similar this state is to the state at the beginning of each of the blocks that can be drawn.

4.2.2.6. Block bootstrap concepts employed in the model

In the current research the following methods, techniques or concepts mentioned above are used to model the distributions of returns:

1) Drawing with replacement from the empirical distribution;
2) Blocks (of monthly returns data);
3) Overlap: The block of monthly returns will overlap each other, incrementing by one month from one block to the next;
4) Circularity: the data is wrapped in a circle;
5) Random block length: The length of each block will be drawn i.i.d. from the geometric distribution with a stated expected block length; and
6) Probabilistic matching of blocks: In the regime-cognisant version of our model, there is a probabilistic matching of consecutive blocks by referring to the multivariate distance between the regime identifying variables at the end of the last drawn block and the beginning of the next drawn block.

In all cases the drawing with replacement will be applied to the historical monthly returns of the portfolio (i.e. asset allocation) under investigation. To be more specific, for any given portfolio, its weights will be applied to the each monthly set of asset class returns to find the portfolio return for each month, and resampling takes place from this set of returns.

Blocks

The logic for resampling blocks instead of single returns in this setting should by now be apparent. By using blocks of returns instead of single months, dependencies over time that may exist in asset class returns are partially preserved. As already mentioned, stock returns have long been thought to exhibit momentum and mean reversion effects, to name but two phenomena.

Overlapping blocks

By overlapping blocks, the number of data points is in a sense increased and the data leveraged to a maximum degree. There are far more rolling periods than there are mutually exclusive blocks, and there is no reason why a block of say 12 months starting in one point in time is more relevant than another 12-month block starting at a different time. By employing different stating points, we are resampling from a larger and richer sampling space. As per our discussion above, overlapping blocks (in combination with circularity) will also mean that the distribution of our resampled data will be the same for each point in time (at least in the regime-ignorant version of our model).

Random block length

As discussed, the block length is one of a very small number of parameters, if not the only parameter, that the practitioner of a block bootstrap method needs to decide on. However, there are very few methods available to help make this decision. Also, due to the fact that the data generating process (in this case the joint distribution of asset classes across time) is generally unknown, the optimal block length can never be known with certainty. Further, even if the optimal block length could be identified, it may not be consistent across all risk measures under investigation. In fact, not only could it depend on the choice of risk measure, but, as we will see, in our current context it would likely also depend on the asset composition of the portfolio under investigation – the optimal block length for a cash-heavy portfolio may be different from the optimal block length for an equity-heavy portfolio due to the potentially different cycles and characteristics such as momentum, mean reversion and other effects of these respective asset classes. It thus makes sense to employ the stationary block bootstrap's random block length, as it diminishes the importance of the choice of block length.

Circularity

Circularity of the stationary bootstrap ensures that the mean of the bootstrapped samples are automatically centred on the data by ensuring that all monthly returns appear in the same number of blocks. The simulated returns are thus as a result also centred on the data. As already mentioned, this ensures that the expected return of the portfolio under investigation is not inadvertently altered.

Probabilistic matching of blocks

In order to incorporate regimes, blocks are resampled not with equal probability, but rather with reference to the prevailing regime (as decided by macroeconomic and market variables to be introduced later), thereby probabilistically matching consecutively resampled blocks of returns. This will become clearer when the full models are defined in Section 4.3.

4.2.3. Our regimes-classification methodology

The economic data introduced in Section 4.1 are the inputs of our regime "classification" methodology. However, each period is not classified as belonging to one of a small number regimes or clusters, as is typically the case. Each month-end is instead considered a state or regime, regardless of how near or far it is to other points in time. Instead of classification, economic and market data will be used to quantify the multivariate distances between the different states. These distances will determine resampling probabilities (in our empirical distribution) conditional on the state the simulation finds itself in. The end result is to resample form an empirical distribution that is conditional on the current state.

For example, when the distribution of returns is modelled conditional on the state being 31 March 2017, the multivariate distance (as measured on our regime-identification variables) of every state relative to the state at 31 March 2017 determines its resampling probability.

The particulars of our regime-scheme, and the various regime-changing processes investigated, will be explicated and formalised under Section 4.3.3, where the full model is specified.

To summarise:

- Every month-end is considered a state;
- Every state month-end is defined by the regime-identification variables;
- Thus, the multivariate distance between any two states can be calculated;
- For any given prevailing state, the closer the other states are in multivariate distance to this state, the higher the probability of resampling the asset class returns belonging to that state; and
- Thus, the distribution of returns for any given state will have a weighting to the returns belonging to all states, but higher weighting to more similar states.

Put another way, our methodology reconciles two incontrovertible, if somewhat paradoxical facts:

(i) Each point in time is in fact unique; and
(ii) The returns of asset classed during any one point in time contain relevant information about the returns during any other point in time – it is merely a matter of, relatively to other points in time*, how much* information. The returns are, after all, generated by the same asset or collection of assets.

There are several gains relative to the typical approach of classifying each point in history as belonging to one of a small number of states or clusters:

- The risk of misclassification is greatly diminished;
- It avoids the need to specify the correct number of regimes or states, which is often based on ad hoc methods, arbitrarily chosen, or limited by the numbers of data points belonging to each regime;
- As already mentioned, every point in time is allowed to draw information about any other point in time, and this information is weighted by how similar the points in time are to each other;
- It is arguably more realistic in its characterisation of each point in time as being unique; and
- The returns distributions will be more robust, because it is a function of returns across all periods, rather than some subset of points belonging to the cluster or regime.

There are two main downsides to this regimes classification methodology:

- There is a loss of interpretability, and the model is more of a so-called "black box". However, this can to an extent be overcome by studying either the multivariate distance or the sampling probability based on this distance of the actual current state to the states during previous points in time. It can also be argued that the loss in interpretability is necessary to achieve a realistic and useful model. We will return to the black box issue after we have expounded on the procedure that produces these distances and probabilities below; and
- There is no clear method for converting multivariate distance to a resampling probability. It is thus not clear whether the resultant weighting given to each state (given the prevailing state) is appropriate.

4.2.3.1. The choice of regime identification variables

We have already listed the economic variables used to measure the multivariate distance between different points in time in Section 4.1. The choice of variables was to a large degree inspired by the industry research of Munro & Silberman (2008) and the academic work of Kaya & Lee (2010).

Munro & Silberman confine themselves to the South African setting. They set out to find the optimal allocation to local asset classes for four economic regimes defined relative to the level and direction of local GDP and inflation:

Economic environments		Inflation	
		Accelerating	Decelerating
GDP Growth	Rising	Expansion	Correction
	Declining	Stagflation	Contraction

Table 10: Munro & Silberman regime classification (excerpt from Munro & Silberman, 2008:1)

Each point in time is classified as belonging to one of these regimes by simple visual inspection of a graph with the two variables. After this classification into mutually exclusive categories, the optimal portfolio for each economic environment is found using only the returns data over the periods deemed to belong that regime.

The regime-classification we will use (detailed in Section 4.3.3) will be different in several important ways to the methodology of Munro and Silberman. Our scheme does not have a small number of mutually exclusive states; instead each point in time is considered a state, and the states will be related to each other by multivariate distance. However, the notion that an economy at any point in time can be quite concisely described with reference to the level and direction of inflation and growth experienced at any point in time, is incorporated into our model. As a result we use both y-o-y GDP growth and inflation figures to indicate the current level of these measures, as well as their 12-month changes, to gauge the direction and rate of change in their levels.

We then also generalise this idea to two economies: the local economy and the US economy. This is a logical step considering that global asset classes are included in our asset allocation optimisations.

US GDP and inflation data are used rather than global data, simply due to the availability of good quality US data extending sufficiently far back in history. While global GDP and inflation figures have been available more recently, older data is more difficult to come by. While not ideal, the size and importance on a global stage of the US economy means it is a reasonable rough proxy for the global economy.

The approaches of Kaya & Lee (2010) and Turner & Han (2010) to identify regimes both have similar elements to our approach.

Kaya & Lee also calculate the multivariate distances between every point in time. They argue for the inclusion of market variables for regime identification purposes. They include S&P 500 and long-term government bonds as two of their regime identifying variables, arguing that if capital markets are forward-looking pricing mechanisms, they necessarily incorporate market expectations regarding the economy and should be useful in the classification of regimes. To incorporate the view of the market with respect to valuations of equity markets, the 12-month return of the ALSI and S&P 500 and the real earnings yields of these two benchmarks are also included. Lastly, the 12-month change in USD/ZAR exchange rate is also included, as this could similarly help pinpoint regimes and describe the market's views on the SA economy versus the US economy.

The work of Turner & Han (2010) is similar to the work in this thesis in the sense that it also employs all observations, but weighted according to distance from the current regime in PCA space.

As discussed above, our regime-classification methodology will employ the notion of multivariate distance (between the values of our regime-identifying variables at different points in time) in order to quantify how dissimilar the state of the markets and economy at different points in time are to each other. Ultimately these multivariate distances will determine the probability (conditional on the current state) of resampling the asset class returns belonging to any other state. In the next section the choice of multivariate distance, namely the Mahalanobis distance, is discussed.

4.2.3.2. Overview of the Mahalanobis distance

The Mahalanobis distance is a multivariate measure of distance between a point and a distribution, or between points, assuming they are from the same underlying distribution. It was first defined by Mahalanobis in 1927 in the context of classifying human skulls as belonging to one of several races on the basis of various measures of its dimensions (Mahalanobis, 1927).

The Mahalanobis distance between two points $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as:

$$(\boldsymbol{x}-\boldsymbol{y})'\boldsymbol{C}^{-\mathbf{1}}(\boldsymbol{x}-\boldsymbol{y}) \tag{60}$$

where $\mathbf{C}$ is the covariance matrix of the underlying distribution, typically estimated by the sample covariance matrix. In the context of two *points*, the Mahalanobis distance answers the following question: Given that two points were produced by the same underlying multivariate distribution, how far are they from each other?"

Similarly, the distance between a point $\boldsymbol{x}$ and any *set* of observations is defined as:

$$(\boldsymbol{x}-\boldsymbol{u})'\boldsymbol{C}^{-\mathbf{1}}(\boldsymbol{x}-\boldsymbol{u}) \tag{61}$$

where $\boldsymbol{u}$ is the mean of the set of observations.

The above expression can be thought of as a multidimensional generalisation of the idea of stating distance in terms of standard deviations from the mean. As it takes into account the covariance between variables, it by definition scales variables by their standard deviations and adjusts distance to take into account the correlations between variables. For example, if two variables are exactly correlated, the Euclidian distance will essentially double-count the effect of those variables, whereas the Mahalanobis distance would not. If the covariance matrix is the identity matrix, it is clear from the formula for Mahalanobis distance that it simplifies to the Euclidean distance.

The Mahalanobis distance is intimately related to the principal components of the variables. If one performs principal components on scaled variables, the Mahalanobis distance between any two points is equal to the Euclidean distance on the axes defined by the full set of principal components. In other words, the Mahalanobis distance has the very attractive property that it identifies and makes use of the axes that best explain the variance found in the scaled data (Brereton, 2015).

4.2.3.3. Mahalanobis distance in our current problem setting

As is evident in Table 11 below, there is a relatively sizable correlation between some of our variables:

	SA Real GDP growth (y-o-y)	US GDP growth (y-o-y)	12m change in SA Real GDP growth (y-o-y)	12m change in US GDP growth (y-o-y)	SA CPI y-o-y	USCPI	12m change in SACPI (y-o-y)	12m change in US CPI (y-o-y)	ALSI price (y-o-y)	S&P500 price change (y-o-y)	ALSI real EY	S&P500 real earnings	ZAR/ USD (12m appreciation)
SA Real GDP growth (y-o-y)	1.00	0.29	0.64	0.18	-0.12	0.36	0.29	0.31	0.04	-0.27	0.15	-0.30	0.02
US GDP growth (y-o-y)	0.29	1.00	0.26	0.44	0.34	0.68	0.14	0.23	0.13	-0.01	0.16	-0.67	-0.07
12m change in SA Real GDP growth (y-o-y)	0.64	0.26	1.00	0.43	-0.08	0.14	0.10	0.43	0.12	-0.14	-0.02	-0.22	0.11
12m change in US GDP growth (y-o-y)	0.18	0.44	0.43	1.00	-0.23	-0.04	-0.15	0.18	-0.08	0.10	0.21	-0.21	0.11
SA CPI y-o-y	-0.12	0.34	-0.08	-0.23	1.00	0.51	0.41	-0.06	0.03	-0.01	-0.61	-0.38	-0.12
USCPI	0.36	0.68	0.14	-0.04	0.51	1.00	0.24	0.36	0.16	-0.17	0.01	-0.62	0.13
12m change in SACPI (y-o-y)	0.29	0.14	0.10	-0.15	0.41	0.24	1.00	0.24	0.06	-0.16	-0.43	-0.28	-0.18
12m change in US CPI (y-o-y)	0.31	0.23	0.43	0.18	-0.06	0.36	0.24	1.00	0.21	-0.16	-0.11	-0.30	0.35
ALSI price (y-o-y)	0.04	0.13	0.12	-0.08	0.03	0.16	0.06	0.21	1.00	0.41	-0.24	-0.14	0.18
S&P500 price change (y-o-y)	-0.27	-0.01	-0.14	0.10	-0.01	-0.17	-0.16	-0.16	0.41	1.00	-0.19	-0.07	0.11
ALSI real EY	0.15	0.16	-0.02	0.21	-0.61	0.01	-0.43	-0.11	-0.24	-0.19	1.00	0.04	0.06
S&P500 real earnings	-0.30	-0.67	-0.22	-0.21	-0.38	-0.62	-0.28	-0.30	-0.14	-0.07	0.04	1.00	0.10
ZAR/ USD (12m appreciation)	0.02	-0.07	0.11	0.11	-0.12	0.13	-0.18	0.35	0.18	0.11	0.06	0.10	1.00

Table 11: Correlations between regime-identifying variables

From Table 11 one can observe that not only are variables positively correlated to their own first difference, but there is also sizable correlation between the basic variables themselves. If the Euclidean distance were used, the latent, underlying factors causing these correlations would be double-counted relative to other latent factors. However, the Mahalanobis distance very elegantly deals with this problem by weighting each variance-scaled variable's contribution to the overall multivariate distance via the sample correlation matrix.

Since there are 541 month-ends in our analysis, we arrive at a 541x541 matrix in which the upper or lower triangle gives the multivariate distance between any two points in time. For example, the very last row of such a matrix gives the distances relative to the latest point in time (31 March 2017), and therefore can be thought of as being associated with the current (as at time of writing) actual economic state.

4.2.3.4. Converting distances to probabilities

In Section 4.3.3, where the regime-dependent models are specified, it will become evident that for any given state the process finds itself in, a probability is required that the returns belonging to another state will be resampled. The probability can loosely be interpreted as the probability that any given state is the closest to the state currently in, or that the returns belonging to that state is most relevant to the given state.

Every row of the distance matrix can be thought of as being related to a single state, and gives the distance between the state associated with that row, and all the other states.

There are several options available to derive probabilities from the distances. One method is to use a kernel function, for example a normal density function. Each row of distances (i.e. the distances of one specific state to all the other states) is converted with a normal distribution kernel. The mean is zero (consistent with the desired property of a distance of zero resulting in the highest probability). The bandwidth parameter of the kernel in the case of the normal distribution is the standard deviation. The standard deviation for each row is the standard deviation of the distances associated with that state, times a "bandwidth factor". The bandwidth factor in this setting determines how far the resultant probabilities will stray from equally weighted. A bandwidth factor that is very large (say 10), results in essentially equal probabilities, while a bandwidth factor of 1 results in very unequal probabilities.

These kernel outputs will often give very high weights to points that are close in time to each other (as the economy sometimes does not change materially on a month-on-month basis) and virtually no weight to other points in time. Similarly, as is done in Kaya & Lee (2010), we could counteract this effect: points that are within 12 months of each other could be adjusted to be no larger than the maximum weight given to points outside of 12 months of each other (this adjustment is made state by state, i.e. with reference to the maximum of each row).

Lastly, we could normalise the adjusted kernel outputs so that each row adds up to one, giving us our resampling probabilities for every given state.

One problem with this method is that it is possible for very a small number of data points to dominate others, which would come at the cost of robustness and stability of optimal portfolios found. This problem is by no means insurmountable. For any given problem, we can investigate the outcome resampling probabilities and calculate the "effective number of data points" given the state.

The effective number of data points is essentially a weighted average for the number of data points, defined identically to the "portfolio diversification index" of Bradfield, Dugmore & Gopi (2006), with the weights in our case determined by the probability of resampling.

In our research the effective number of data points can be calculated with the following procedure, for any given state i:

$$effective\ datapoints\ (i) = 2\sum_{j=1}^{N} p(j|i)\, r_j^{(i)} - 1 \qquad (62)$$

where $p(j|i)$ is the probability of resampling state j given the process is in state i, and $r_j^{(i)}$ is the rank of the j^{th} probability among all probabilities associated with state i. N here is total number of states or month-ends (541 in our analysis).

If an equal weighting is given to every point, the effective number of data points is just the actual number of data points, namely 541, as one would expect of such a measure.

If the effective number of data points is too small, there is an over-reliance on just a small number of returns in our resultant returns distribution. We can spread the probability around more by increasing the bandwidth as required.

However, this does require more supervision and time. In Section 5.2 a large number of optimal portfolios are derived and backtested by varying the end point in terms of time of the analysis. To avoid having to investigate the effective number of data points for each point in time, we opted for a simpler weighting scheme than the more sensitive kernel method discussed above: an exponential weighting based on the percentile rank of the distances for a given state. Under this scheme the probability of resampling any one state j given the state i, is

$$p(i,j,P) = \frac{e^{P(r_j^{(i)}-1)/N}}{\sum_{j=1}^{T} e^{P(r_j^{(i)}-1)/N}} \qquad (63)$$

where P is defined as the "tactical pressure".

As the percentile rank will, for a fixed number of periods, always result in exactly the same series of numbers, it is known *a priori* how much weight the highest weighted to lowest weighted observation will have, thereby ensuring the effective number of data points is known, regardless of the values of the regime-identifying data.

The tactical pressure will determine how much skewing of the resampling probability from equally weighted will take place and therefore generally result in larger tactical tilts from the strategic allocation (the outcome if equal probabilities are used). This parameter plays a similar (but inverted) role as the bandwidth parameter in the kernel method discussed above. It can be viewed is a lever that the practitioner could use as an expression of how aggressive he chooses to be, tactically, at any point in time. Table 12 and Table 13 below show the ratio of the data point with highest resampling probability versus different percentiles, and the cumulative resampling probability for different percentiles, respectively:

		probability of selected percentiles (as ratio vs highest probability)						
pressure	eff datapoints	0.0%	5.0%	10.0%	25.0%	50.0%	75.0%	100.0%
0.00	541	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	452	1.00	1.05	1.10	1.28	1.65	2.11	2.71
1.35	423	1.00	1.07	1.14	1.40	1.96	2.75	3.84
1.80	387	1.00	1.09	1.20	1.57	2.46	3.85	6.01
2.50	336	1.00	1.13	1.28	1.87	3.48	6.50	12.07
3.50	275	1.00	1.19	1.42	2.39	5.74	13.74	32.69
4.50	228	1.00	1.25	1.57	3.07	9.45	29.04	88.53
5.50	192	1.00	1.32	1.73	3.95	15.56	61.40	239.77
10.00	108	1.00	1.65	2.71	12.13	147.05	1783.15	21227.05

Table 12: Ratio of maximum probability to the probabilities associated with selected percentiles

		cumulative probability of selected percentiles of probabilities						
pressure	eff datapoints	0.0%	5.0%	10.0%	25.0%	50.0%	75.0%	100.0%
0.00	541	0.2%	5.2%	10.2%	25.1%	50.1%	75.0%	100.0%
1.00	452	0.3%	8.0%	15.3%	35.2%	62.3%	83.5%	100.0%
1.35	423	0.3%	9.1%	17.3%	38.8%	66.3%	86.0%	100.0%
1.80	387	0.4%	10.7%	20.0%	43.6%	71.2%	88.8%	100.0%
2.50	336	0.5%	13.2%	24.5%	50.8%	77.8%	92.3%	100.0%
3.50	275	0.7%	17.1%	30.9%	60.3%	85.3%	95.7%	100.0%
4.50	228	0.8%	21.0%	37.1%	68.5%	90.5%	97.7%	100.0%
5.50	192	1.0%	24.9%	43.0%	75.2%	94.0%	98.8%	100.0%
10.00	108	1.8%	40.4%	63.8%	91.9%	99.3%	99.9%	100.0%

Table 13: Cumulative probabilities of percentiles

Table 12 shows that for a pressure of 10, the point with the highest probability to be resampled will be resampled 1.65 times as often as the data point that corresponds with the 5th percentile, and 1783 times as often as the lowest ranked data point.

Table 13 shows that for a pressure of 10, the 25% of points closest to the current point will be given 92% of all resampling weight.

As already mentioned, the above tables are completely independent of the underlying data in the dataset, which means that the effective number of data points is known in advance.

However, this comes at a loss: the resampling probabilities are no longer sensitive to the *relative* Mahalanobis distances between points – all that matters is the *ranking* of the distance. At first glance, this seems like a material disadvantage. However, in the absence of a known, direct relationship between asset class returns and the regime identifying variables, a scheme that is too sensitive to the relative magnitudes of distances may, in fact, be inappropriate. If the relationship is unknown, it is preferable to use a smoothly changing weighting system where the progression in resampling probability changes gradually from the closest state to the farthest state, as depicted in Figure 20:

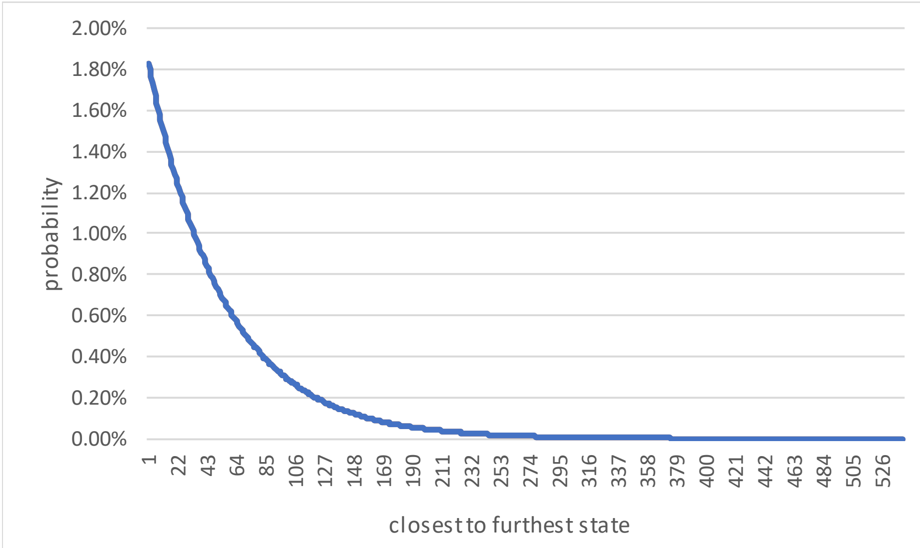

Figure 20: The highest to lowest probabilities of resampling (tactical pressure = 10).

> *These ordered probabilities are invariant to the prevailing regime (though the points in time each of them represents is dependent on regime). The approximate resampling probability for the state that is closest to the current state is approximately 1.8%. This graph will look exactly the same for any fixed pressure parameter, regardless of the current state or the choice of distance measure; however, the ranking of the closest to furthest state will change, thus the numbers on the horizontal will refer to different states if the current state is changed.*

Nevertheless, for any two given points, the probability of resampling will be higher if the distance to the current regime is lower (and vice versa) as the resampling probabilities are sensitive to the relative ranking of the distances to prevailing regime.

The full process of calculating the resampling probabilities for the various states can thus be summarised as:

1) Find the Mahalanobis distances between the regime-identifying variables as at every point in time (in our case, there are 541 month ends, and thus 541x540 such relationships (e.g. the upper triangle of a symmetric 541x541 matrix).
2) For each row in this matrix: rank each entry from lowest to highest and divide rank by number of entries to arrive at percentile rank, and apply expression (63) to each rank in this row.
3) The rows of the resultant matrix now each represents the sampling probabilities conditional on the regimes they represent.

4.3. The models

4.3.1. The parameters that define our problem settings

In Chapter 5 the optimal portfolios for specific problem settings are discussed. These problem setting will defined by the following parameters:

- A required real return;
- A risk criterion to minimise;
- An investment horizon ($\boldsymbol{T}$) – the period over which the required return should be achieved, and over which the risk criterion is (usually) evaluated; and
- An expected block length ($\boldsymbol{L}$)[18]

The effect of each of these parameters on the optimal portfolio is isolated and examined in some detail in Chapter 5.

As already mentioned, the returns of any one portfolio are modelled with a slight variation of the empirical distribution of its historical returns. The basic outline of our returns modelling can be summarised as follows:

- First, find the $\boldsymbol{N}$ monthly returns for the portfolio under investigation by applying the dot product between its asset weights and the monthly returns of the asset classes (if there are $\boldsymbol{N} = 541$ months of data then there are $\boldsymbol{N}$ dot products).
- The resultant monthly returns are wrapped in a circle, with the last monthly return followed by the first. That means that there are always $\boldsymbol{N}$ selectable blocks, irrespective of the block length.
- Then randomly resample one block after the next, with the block length of each and every block distributed (i.i.d.) geometrically, until we have a sufficient number of blocks to cover the investment horizon.
- Next calculate the return and the risk criterion for this one iteration.
- Repeat above steps a large number of times, and then find the average return and average risk criterion.
- In the first model (ignoring regimes), the resampling probability of blocks are always equal. For the regime-cognisant models, depending on the prevailing regime, some blocks will be resampled with higher probably than others.

4.3.2. Regime-ignorant model

This section describes the resampling scheme when we ignore regimes and resample blocks with equal probability, with an expected block length of $\boldsymbol{L}$ for an investment horizon of $\boldsymbol{T}$ months:

[18] Strictly speaking, the expected block length does not form part of the problem setting but is instead a nuisance parameter of the returns distributions model.

- ❖ ***For*** *i=1 to the total number of bootstrap iterations (say 5000):*
  - ➢ ***While*** *the cumulative length of resampled blocks is less than the horizon length* ***T***
    - ▪ *Sample a block length from geometric(**L**) independently*
    - ▪ *Sample a start position of the block in the dataset (from discrete uniform distribution(1,**N**))*

      #The above two steps in combination define one selected block
  - ➢ ***End while***
  - ➢ *Discard the extra months of returns over and above* ***T****.*
  - ➢ *Calculate the relevant risk measure and real return over the bootstrapped*

Algorithm 1: The joint asset class distribution model that ignores regimes

The most important characteristics of this model are:

- It is informed by and highly reactive to the historical returns of each and every portfolio;
- The dependencies that exist (including non-linear dependencies, such as volatility) over time are preserved within blocks;
- The shape of the distribution found in the data is conserved; and
- Blocks are treated exactly the same: they are resampled with equal probability, no block is given higher importance (or deemed to be more relevant) than any other blocks.

4.3.3. Regime-cognisant models

As already explained in Section 4.2.3, here regimes are incorporated by resampling blocks with unequal probabilities that are based on the multivariate distance between the values of the economic and market regime identifying variables. Ultimately we generate empirical distributions conditional on the current regime.

The number of possible states is therefore the number of months of data in our analysis, namely 541. The state can be understood as an economic and market regime, although each state is unique and treated as a regime, as opposed to the typical regime approach, which classifies all periods into typically between two and six regimes.

How states change or do not change between resampled blocks lead to three distinct regime-cognisant models, each with its own real-world interpretation.

4.3.3.1. The three different state changing mechanisms

In these descriptions below, we will refer to

- the *current state*: the multivariate values of the regime-identifying variables as at the date of the analysis (in this research, the current state is as at 31 March 2017, as that was the latest available data at the time of running of the models); and
- the *prevailing state* (in the bootstrap simulation): in two cases, the state will change from one resampled block to the next. At any one point in time, the state the algorithm finds itself in is referred to as the prevailing state.

Three distinct possible state changing *mechanisms* are briefly described.

4.3.3.1.1. Regime-scheme 1

The prevailing state is always (statically) the current state of the regime-identifying variables.

The state is fixed to the latest macroeconomic/market state, i.e. *the prevailing state* is always the *current state*. In this regime setup, we are always resampling blocks with respect to this static state. In other words, the probability of resampling any block will always depend on how far it is in terms of Mahalanobis distance from this static current state. The structure of this model is identical to the regime-ignorant version defined in Section 4.3.2, except that the resampling probabilities are unequal (but still static).

4.3.3.1.2. Regime-scheme 2

At the start of each iteration, the prevailing state is set to the current state, but then changes within an iteration to the state at the end of the resampled block.

The first block is again resampled with respect to the actual current regime (31 March 2017). The state then changes to the state at the end of the previously resampled block within each iteration. Once we have a sufficient number of returns (to cover the investment horizon at hand), *the state returns to the actual current state for each new iteration*.

- ***For*** *i=1 to the total number of bootstrap iterations (say 5000):*
  - ♦ *Set state to current state (31 March 2017)*
  - ♦ ***While*** *the cumulative length of resampled blocks is less than the horizon length **T***
    - ➢ *Sample a block length from geometric(**L**) independently*
    - ➢ *Calculate resampling probabilities based on the prevailing state*
    - ➢ *Sample a start position of the block in the dataset using above resampling probabilities*
    - ➢ *Set the state to the state as at the end of the resampled block*
  - ♦ ***End while***
  - ♦ *If we have more than **T** months of returns, discard the extra months of*

Algorithm 2: Joint asset class returns distribution model conditional on regimes (regime-model 2; state changes to the end of the last resampled block)

4.3.3.1.3. Regime-scheme 3

At the start of each iteration, the prevailing state is set to the current state. However, within each iteration, the prevailing state then changes to the end of the first generated block and then always moves chronologically forward in time from there by each successive generated block.

In other words, the first block in each iteration determines where the prevailing state jumps to, and from there the states change as they did in the actual chronological history of returns. Again, the data wraps around itself, so if 31 March 2017 is reached, the next state is the earliest data included, i.e. 31 March 1972. The state returns to the current state for each new iteration.

This will become clearer with the full formalisation of the algorithm:

- ***For*** *i=1 to the total number of bootstrap iterations (say 5000):*
  - ♦ *Set state to current state (31 March 2017)*
  - ♦ *Calculate resampling probabilities based on the current state*
  - ♦ *Sample using above resampling probabilities a start position*
  - ♦ *Sample a block length from geometric(**L**) independently*
  - ♦ *Set the state to the state as at the end of the resampled block*
  - ♦ ***While*** *the cumulative length of resampled blocks is less than the horizon length **T***
    - *Sample a block length from geometric(**L**) independently*
    - *Calculate resampling probabilities based on the prevailing state*
    - *Sample using above resampling probabilities a start position of the block in the dataset*
    - *Move the state forward by the last generate block length*
  - ♦ ***End while***
  - ♦ *If we have more than **T** months of returns, discard the extra months of returns.*

Algorithm 3: Joint asset class returns distribution model conditional on regimes (regime-model 3; state changes chronologically after the first resampled block)

4.3.3.2. Discussion of the three schemes

Regime-scheme 1

The first scheme is the simplest, keeping the state fixed. In practice it would usually be the state defined by the latest regime-identifying macroeconomic and market variables. Implicit in the first method is the assumption that attempting to predict future states beyond the simple estimate of the current state is futile: the current state is considered the best predictor of future states across the entirety of the investment horizon. As discussed in the literature review of regimes, regimes are generally highly persistent. However, as is the case for all our regime mechanisms (as well as the regime-independent model), regime-changes within blocks are implicitly modelled: whatever regime changes occurred within blocks are preserved within blocks.

A strength of this scheme is its simplicity and interpretability: the optimal portfolios derived will be relevant to the current, static state. A weakness is that the longer the investment horizon, the less and less realistic is the assumption of an unchanging state. Further, the current state is heavily emphasised. If the underlying data is spurious or highly idiosyncratic, the resultant distribution may be extreme or unrealistic. However, predicting future macroeconomic and market states is such a difficult enterprise, that arguably the current state is our best predictor of future states, especially for shorter investment horizons.

Regime-scheme 2

Both the second and third methods allow states to change over time, but do so in very different ways.

The second method starts each iteration at the state as at the date of the analysis (31 March 2017), and then changes states after each and every block is resampled to the state as at the end of said resampled block, with the probability only depending on the current state. The state thus changes within iterations according to a second-order Markov process, where the probabilities are based on the multivariate distance between the states.

The implication or assumption is that the economic and market variables are simply more likely, after each block, to transition to similar states than to dissimilar states. A strength of this approach, relative to the first method, is that we place less emphasis on just one state (i.e. the 31 March 2017 state). In the real world, the state of the economy does change from month to month. Even if we cannot necessarily capture exactly the state changing mechanism, the mere fact that we allow states to change to states in close proximity, may improve the robustness of the solution.

When we implement our regimes-based resampling schemes, a relatively long expected block length of 12 months, and relative short investment horizon of 36 months are used. In other words, there will on average, roughly speaking, be only two state changes – at the end of 12 months and then again at the end of 24 months. This naturally limits the effect of any potentially unrealistic regime transitions while still reducing reliance on the regime as at the time of analysis.

A weakness of this approach is that it is not clear that the state changes in a realistic manner across time. Does the macro-economic state jump too abruptly across a single iteration if one considers the underlying regime identifying variables individually? Are we properly modelling the cyclical nature of macroeconomic and markets? This is not investigated further. However, an opportunity for further research would be to describe the implied Markov process, find the long run probability of being in any one state, and to investigate the realism of these outcomes.

Regime-scheme 3

The third method jumps to a new state at the beginning of each iteration. Thereafter, state changes are guaranteed to happen in a realistic manner (apart from the one discontinuity found at the end of the sample where data wraps around back to the beginning), as we are then in essence using historical state time paths after the initial jump. If the strength of this scheme is that the state changes are guaranteed to occur realistically, the weakness is that we are limited in to a single transition path, and thus very dependent on the historical evolution of regimes over the period of our sample.

The three regime-schemes we have defined can be compared to the typical Markov-switching model pdf given by expression 25 on page 26.

The most important difference to the typical approach is that instead of imposing a normal density (or some other family of distributions) for the conditional densities, we employ conditional empirical distributions, allowing for the possibility that different states have distinct distributional shapes associated with them. Also, in our regime-schemes, both the transition probabilities and the empirical distributions may depend on the probabilities calculated with expression 63 on page 77.

Regime-scheme 2 perhaps has the most in common with the typical Markov-switching model, the main difference being the use of empirical distributions instead of normal densities for the conditionals.

In the case of regime-scheme 1, the state does not change at all, thus the summation of expression 25 falls away and there is only a single conditional empirical distribution with probability of 1.

Regime-scheme 3 randomly selects a first state based on the transition probabilities of expression 63, but thereafter the only source of randomness in state jumps within an iteration is the random block length, otherwise state changes follow their historical chronological path.

Generalised empirical distribution

All of these regime mechanisms could be considered generalisations of the empirical distribution. Instead of giving every data point equal weighting, we weight them according to their relevance to the current regime. Why do we do this? It is likely that the expected returns, variability of returns, and the stylised facts and other aspects of the returns distributions of asset classes depend on the regime identifying variables. In each case, we are *conditioning for the current state,* ultimately finding the empirical distribution conditional on the current state.

The resultant weighted empirical distribution belonging to each state will be reactive to the unique characteristics of the actual joint distributions of asset classes, taking into account regimes or states to the extent that these effects are present in the returns data. A clear advantage of the nonparametric approach thus becomes apparent: instead of imposing a single distribution with a specific shape to all regimes, we instead use the versatile empirical distribution to let the returns data themselves inform the unique shape of the distribution for any given state.

4.3.3.3. The tactical pressure factor

In the broader context of this research, the tactical pressure factor defined in Section 4.2.3.4 can now be more fully understood to be a tactical or regime asset allocation lever: the higher the pressure, the more unequal the weights given to data points, the more aggressive will the resultant optimal portfolio be relative to the optimal portfolio that gives all data points the same weight (which could be considered the strategic or long-term asset allocation). If the pressure is sufficiently close to zero, the optimal portfolio ignorant of regimes is retrieved. It is therefore one of the few important parameters of this research (alongside expected block length).

In practice the pressure factor would be where the asset manager decides on how aggressively he wants to tilt his active portfolio at any point in time. How aggressive an asset manager chooses to be depends on his conviction and risk appetite. The selection of the pressure can be an iterative process of experimenting with its level until the tilts are feasible or the resulting historical tracking error is acceptable.

The choice of pressure could also depend on the data itself. For example, pressure may be proportional in some sense to how far away the regime identifying variables are from its centroid in multivariate distance terms. Another possible avenue is to iteratively solve for this parameter by performing cross-validation with subsequent performance and inferring the optimal constant value across time. Deriving and discussing such a scheme falls outside the scope of this research, but is a definite possible avenue for further research.

For the purposes of reporting results for the regimes models, a mostly arbitrary pressure of 10 is employed. This number was chosen simply by observing the resulting resampling probabilities, which were adjudged to be likely to result in relatively robust distributions while still having a material effect on the weighting scheme and hence would, all considered, result in materially tilted yet robust optimal portfolios.

4.3.3.4. Overcoming the black box nature of the regimes classification method

Typical regime methods would classify all data into a small number of mutually exclusive and easily interpretable clusters. The current state at the date of the analysis is classified as belong to one of the clusters or states. Often these states can be understood as intuitively good or bad states (in the economic or market sense), or some combination of high/low growth and high low/inflation.

Our scheme is not so simple and interpretable. However, by investigating the multivariate distances of the current actual state to previous states, we could get a sense of the current state and how it relates to earlier states. Equivalently, by examining the resampling probabilities based on these distances, the user of the model will begin to understand how much weight is given to each state in history in the resampling procedure.

For example, for the state as at 31 March 2017 and a tactical pressure parameter of 10, the resampling probabilities for every other state are shown in Figure 21:

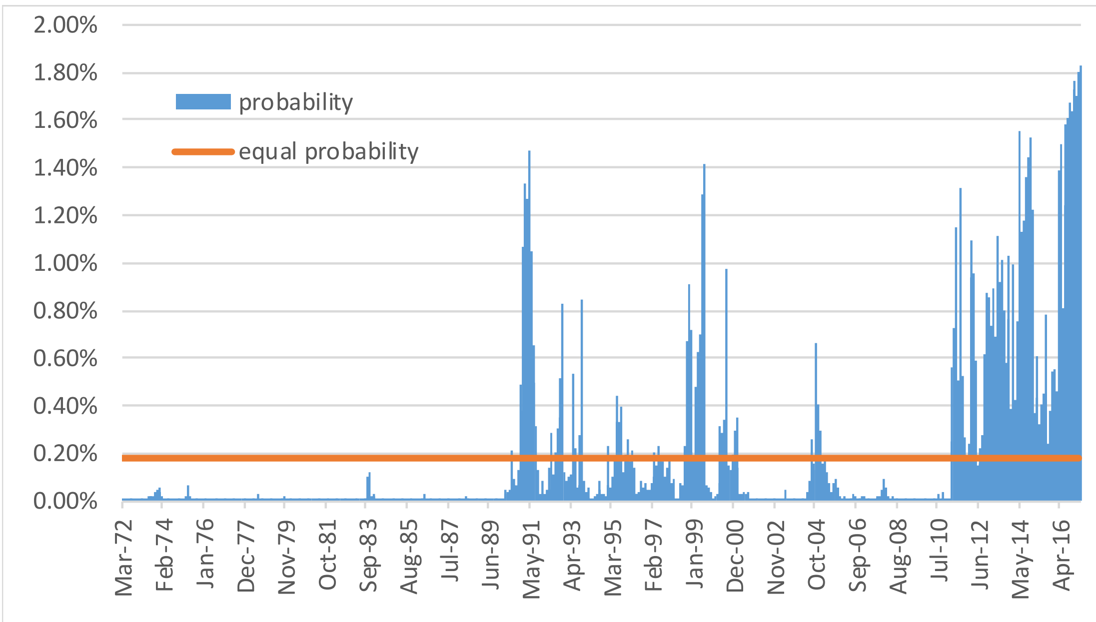

Figure 21: Probabilities associated with state as at time of analysis (31 March 2017) and tactical pressure of 10

In the case of regime-scheme 1 (where the state remains constant), the information in Figure 21 helps us to understand very directly how each historical period will inform our conditional empirical distribution for returns and the resultant optimal portfolios. If the probability is below the horizontal line for a given date, the resampling probability for the returns belonging to that month is less than the equally weighted (1/N) probability, and vice versa.

For regime-scheme two, the state changes according to a Markovian process. For any fixed block length and horizon, we can calculate analytically the probability of being in any given state after every block is selected (i.e. state change takes place). However, in all our optimisations the block length will be randomly generated after each block is selected, which complicates the mathematics somewhat, but it would be simple to simulate the probability of being in any state at any point in time.

A similar simulation can be performed for regime-scheme 3 to better understand how we are resampling from history. While these simulations are not performed in this research, we note that this could be an additional analysis performed by a practitioner to gain more insight into the conditional empirical distributions belonging to the current state.

4.3.4. The choice of expected block length

The choice of block length is, in essence, a bias-variance trade-off: if the block length is equal to the investment horizon, our estimates for the joint distribution will be unbiased. If the block length is one month, the variance of the estimate for the underlying distribution will be minimised, but will generally be biased. For example, consider our implicit estimate for the variance of the returns over the investment horizon of any portfolio under consideration: if the block length is equal to the investment horizon, our estimate is simply the historical variance of this portfolio over a periodicity equal to the investment horizon. However, if the block length is shorter than the investment horizon, bias is introduced due to the time-dependencies of asset returns discussed in Section 4.1.4.2 (for example, in the case of SA equities, bias would arguably be induced due to either short-term momentum or long-term mean reversion, depending on the length of the investment horizon).

One could define the optimal block length as that length which minimises the standard errors of the estimates of the underlying distribution. In the previous section, we showed how the annualised standard deviation of SA equities depended on the periodicity. The results there already hint how biased estimates of variance are when the investment horizons and block length are not equal, but tell us nothing of the variance of estimates, and the overall standard errors. To gain perspective on the overall sacrifice in terms of standard error for the gain of less bias when we increase the block length, Figure 22 to Figure 26 depict (for return periodicities of respectively 12, 36, 60, 120 and 240 months) the bootstrapped standard errors[19] for estimates of the variance of the real returns of SA equities, as a function of the choice of block length:

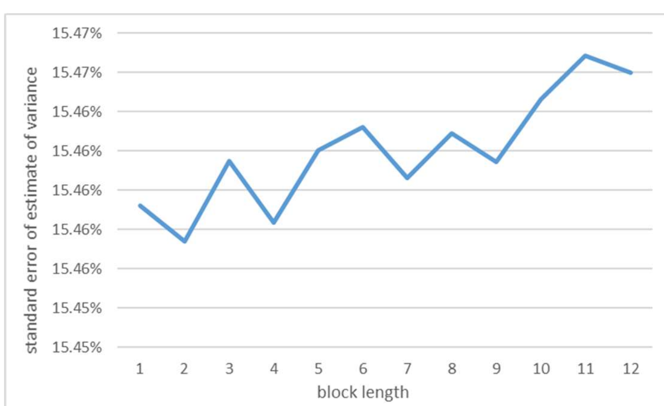

Figure 22: Standard errors for estimating variance of 12-month SA equities returns versus block length

[19] 10 000 bootstrap iterations in a circular block bootstrap with fixed (non-random) block length

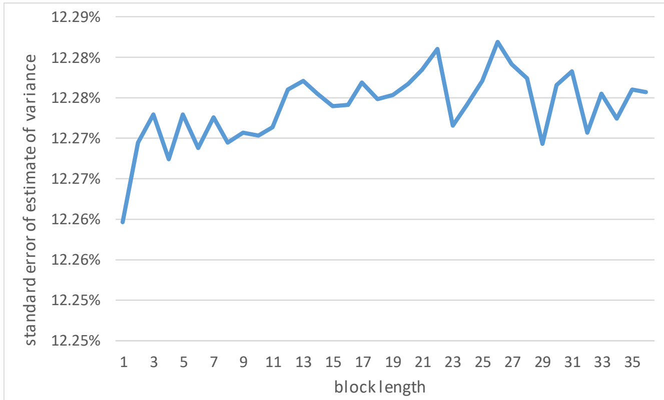

Figure 23 : Standard errors for estimating variance of 36-month SA equities returns versus block length

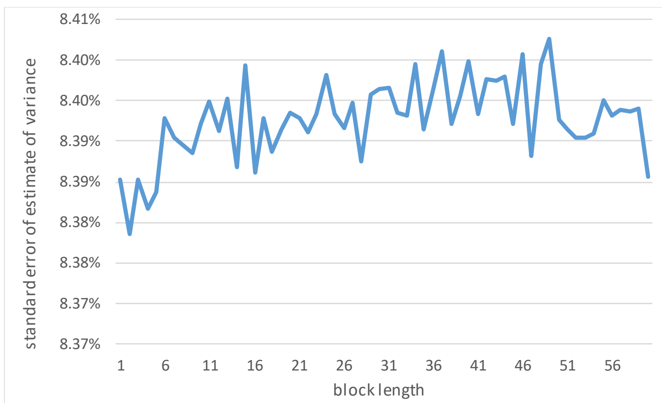

Figure 24: Standard errors for estimating variance of 60-month SA equities returns versus block length

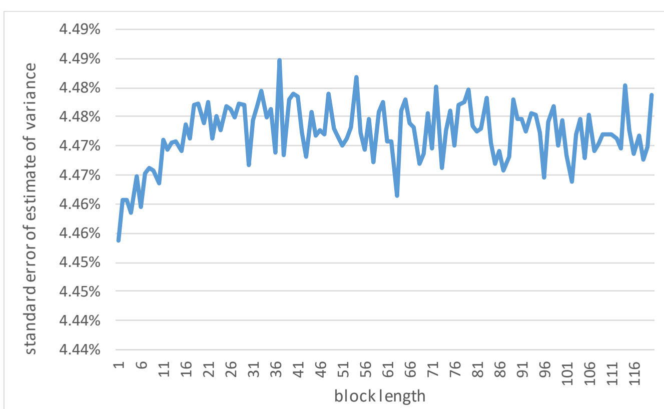

Figure 25: Standard errors for estimating variance of 120-month SA equities returns versus block length

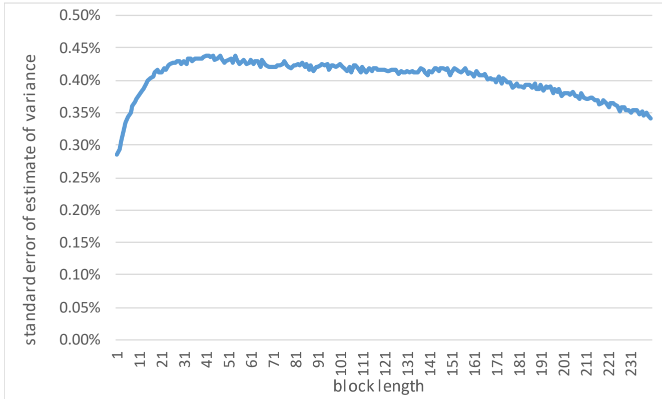

Figure 26: Standard errors for estimating variance of 240-month SA equities returns versus block length

It can be seen in Figure 22 to Figure 26 that the optimal block length as judged by estimates of standard errors of the variance of the real return of SA equities is thus one or two months for all periodicities. However, it is important to note that to gain a reduction in bias, there is very little sacrifice in terms of the overall standard error.

As the variance of each portfolio's returns is dependent on its asset allocation, and the returns of each asset class exhibits its own distinct behaviour, the optimal block length is strictly speaking not generally the same for different portfolios. However, SA equities are generally the asset class with both the largest allocation in optimal portfolios and highest variance, and is therefore the most important asset class in terms of determining the optimal block length. While not shown, a block length of one or two months also minimised the standard error in the case of a portfolio with equal weights to SA equities, SA bonds, SA cash, global equities and global bonds.

Although the importance of block length is diminished by random sampling from a geometric distribution, the choice of expected block length still has material repercussion for the optimal portfolios (as we will show under the results sections in Chapter 5). There is also no known method to determine which block length will result in the most accurate and practically useful optimal portfolios. Therefore, our result section and appendices discuss the optimal portfolios pertaining to three categories of expected block length settings:

- Lowest variance and standard error case: block length of one month;
- Empirical case: block length equal to the investment horizon resulting in the lowest bias in estimate of the variance of SA equities returns; and
- Intermediate case: expected block length somewhere between one month and investment horizon, representing a compromise between bias and variance.

By considering the optimal portfolios for these three different block lengths, we hope to gain some understanding of how estimation errors due to time-dependencies and sampling error may impact our estimates of the optimal portfolios.

Choice of block length for intermediate case

In the intermediate case we will set out to choose the longest expected block length that still results in a relatively well defined distribution. The period of 541 months of our dataset implies the number of available data points for different periodicities tabulated in Table 14:

periodicity	number of mutually exclusive periods in dataset	number of rolling periods
1	541	541
12	45	530
24	23	518
36	15	506
48	11	494
60	9	482

Table 14: Number of definable blocks versus periodicity

Recall that we leverage our data to the maximum degree by using rolling (overlapping) periods, even if we have relatively few independent data points. For example, consider the number of blocks and combinations for an investment horizon of 60 months and 120 months, respectively presented by Table 15 and Table 16:

Horizon (months)	60				
block length (months)	12	24	36	60	84
overlapping blocks	529	517	505	481	457
mutually exlcusive blocks	45	22	15	9	6
required blocks	5	3	2	1	1
possible combinations (overlapping)	3.52E+11	23165219	127765	481	457
possible combinations (mutually exlcusive)	1906884	2024	120	9	6

Table 15: Number of blocks and combinations for horizon length of 60 months (ignoring circularity)

Horizon (months)	120				
block length (months)	12	24	36	60	84
overlapping blocks	529	517	505	481	457
mutually exlcusive blocks	45	22	15	9	6
required blocks	10	5	4	2	2
possible combinations (overlapping)	5.15E+20	3.14E+11	2.74E+09	115921	104653
possible combinations (mutually exlcusive)	2.39E+10	65780	3060	45	21

Table 16: Number of blocks and combinations for horizon length of 120 months (ignoring circularity)

Table 15 shows that for a block length of 60 months, we require nine blocks to cover the investment horizon of 360 months. There are 481 possible ways to select one block out of the 481 overlapping blocks in our dataset (when the order is ignored – all of the risk criteria considered are invariant to the order of blocks). Table 16 shows that this number increases to 115 921 if the horizon doubles to 120 months and the block length remains at 60 months (in reality, we are also employing a variable block length, which further increases the complexity of the distribution).

Our main results for the five asset class case below found the optimal allocations for investment horizons of 12, 36, 60, 84, 120 and 240 months, while varying the block length from 1, 3, 6, 12, 24, 36 to 60 months (as well as the investment horizon itself – see the "empirical case" below). From these options, we ultimately chose, as shown in Table 17 below, an "intermediate" block length for each investment horizon:

horizon	chosen block length	possible combinations (overlapping)	possible combinations (mutually exlcusive)
12	6	143380	4095
36	24	133903	253
60	36	127765	120
84	60	115921	45
120	60	115921	45
240	60	2.26E+09	495

Table 17: Intermediate block lengths for each investment horizon

In the case of each horizon length shown in Table 17, we chose the largest block length available shorter than the investment horizon and that still resulted in a relatively large number of possible combinations, mainly from the overlapping (non-ordered) combinations point of view. In the case of the 36-month investment horizon, the choice of 24 months corresponds with the block length with generally the best performance in out-of-sample backtests conducted with actual asset class returns from 1972 to 2017 (as will be discussed in Section 5.2.3).

4.3.5. Rebalancing and transaction costs

This research assumes a full monthly rebalance every month and ignores all fees and costs, including transaction costs. In practice pooled funds have a 'free" opportunity to rebalance with every cash flow entering or leaving the fund. In the context of funds with daily cash flows and a target allocation, it is arguably more realistic to assume frequent rebalancing (though it does depend on the size and frequency of cash flows, as well as the investment process of the fund manager).

Funds also have several reasons to fully or partially rebalance from time to time even if there are not sufficient cash flows. This may be to align different funds tactically with each other (to be overweight and underweight in same asset classes and to the same degree in, say, low-risk, medium-risk and high-risk balanced funds), or to align them all closer to some ideal tilt relative to the strategic asset allocation. There are several valid investment reasons for this alignment, but also many business reasons.

From an investment point of view, in the absence of rebalancing, the asset class exposure fluctuates over time with some degree of randomness. Depending on which asset class happened to outperform at a certain time, the expected return of the resulting asset allocation may be higher or lower than intended. Assuming a full monthly rebalance helps reduce the effect of this randomness.

The modelling of risk is more internally consistent if the asset class exposure remains constant. A meandering asset allocation results in fluctuating risk exposure. More broadly, as the optimal asset allocation is the very vector we are solving for, it would arguably be incoherent and inconsistent to simply allow it drift in our modelling.

However, it should be noted that the results that will follow are strictly only relevant when there is relatively frequent rebalancing of some nature, and this rebalancing does not result in significant additional transaction costs, such as a fund experiencing frequent ad hoc inflows or outflows. For the most part transaction costs would simply mean that the real return is understated and that, for example, the optimal portfolio is simply for a real return of 1.8% rather than say 2%. However, even if a return adjustment is made, technically the resultant optimal weights would be suboptimal if the transaction costs are significantly different across asset classes.

The real returns shown and discussed will always ignore any fees and costs not inherent to the return proxies employed.

So far in Chapter 4 we have introduced the data as well as various concepts employed to define the model, and, lastly, specified the full models. The following chapter discusses all the optimal portfolios related to a number of different settings each defined by an investment horizon, risk criterion and required real return. Although omitted from the main body of this work, the technical aspects regarding the finding of optimal portfolios can be found in Appendix 7.1.

5. RESULTS

This chapter discusses the optimal portfolios derived in various problem settings. It is divided into two main sections: The optimal portfolios ignoring regimes (5.1), and the optimal portfolios incorporating regimes (5.2).

The empirical nature of our joint distribution, along with the variety of risk criteria employed, pose a significant technical challenge as it calls for a versatile and robust optimisation routine. Standard methods such as quadratic optimisation cannot be employed in our context. We consider the methodology employed to solve for these optimal portfolios to be an important component of this work. However, in the interest of brevity and flow, we omit the technical details from the main body of this work. For those who are interested in these details, Appendix 7.1 introduces genetic algorithms and gives the technical specifications of our bespoke genetic algorithm.

5.1. Results of optimisations ignoring regimes

5.1.1. Defining the problem settings

This section discusses all the optimal portfolios for various problem settings, *ignoring regimes*. As such these optimal allocations are of particular relevance to the question of long-term, strategic asset allocations which, at least in theory, remain static over long periods of time (as mentioned in the introduction to this thesis). In statistical terms, the optimal portfolios derived here are optimal with respect to the *unconditional* joint distribution of asset classes.

The results for the regime-cognisant optimal portfolios will be discussed in Section 5.2.

Five asset case and eight asset case

The main results include only the five asset classes, namely SA equities, SA bonds, SA cash, global equities and global bonds.

Local and global listed property asset classes are excluded from the main case for several reasons. The relatively small sizes of the property asset classes in terms of market capitalisation (SA listed property stocks contribute only in the region of 1% to 2% of the SA equities market), as well as the limited liquidity they offer asset managers, call into question whether they should be considered stand-alone asset classes. The similarity in their return and risk characteristics to their equity counterparts, and the fact that they are listed on the same exchanges, further argue for them to be considered as a category within equities. From an analysis and explication standpoint, their similarity to equities also means that the interpretation of optimal portfolios including both equities and property can be confusing, especially when a large number of optimal portfolios are considered, as will be the case here. Lastly, the return histories of these asset classes are less reliable and potentially less relevant the further we go back in history and the closer we get to when the listed property market was still in its infant stages.

We excluded global cash based on the fact that this asset class in any case generally receives very little allocation in optimal portfolios (typically less than 1%), and again for the sake of interpretation, simplicity and clarity.

However, these three asset classes are included in the eight asset case in Appendix 7.3, and the reader specifically interested in their optimal weights can refer to that section.

Comparisons to MVO portfolios

The results below will occasionally compare each optimal portfolio to two MVO portfolios. Both of these MVO portfolios are on the same efficient frontier (based on monthly data), but have different levels of real return:

(1) "MVO with the required return" or just "MVO" has a real return equal to the required real return of the optimal portfolio in question; and
(2) "MVO with the same average real return" has the same real return as the optimal portfolio that it is being compared to.

The latter is a different portfolio than the former in the case of probability of success, average shortfall, downside deviation and standard deviation over horizon, as these optimal portfolios typically have average real returns *in excess* of the required real return.

Both of these MVO portfolios are of interest, as they represent more traditional options available to investors.

Problem settings

The optimal asset allocations to SA equities, SA bonds, SA cash, global equities and global bonds were found for all of the following problem and parameter settings:

- Required real returns of 2%, 4% and 6%;
- Investment horizons of 12, 36, 60, 84, 120 and 240 months, each with its associated intermediate block length (as defined in Section 4.3.4)*;* and
- The following risk criteria:
    - Average shortfall over investment horizon
    - Downside deviation over investment horizon
    - Probability of success over investment horizon
    - Standard deviation over investment horizon
    - Maximum drawdown
    - Monthly standard deviation (i.e. standard MVO).

The block length is a nuisance parameter (and thus not strictly speaking part of the problem setting) whose specification completes the unconditional returns distributions model. All of the following block lengths were considered (in months): 1, 3, 6, 12, 24, 36, 60 and the empirical case (i.e. a block length fixed to the investment horizon).

However, unless specifically stated otherwise, we employ the intermediate block lengths of Section 4.3.4 in all of our regime-ignorant results below (i.e. each and every investment horizon has one associated expected block length that is primarily considered).

Base case problem

After providing a table with our main results for each of the problem settings, we will, in the interest of brevity and explication, narrow our attention on one specific base case problem, and then one by one vary each of its parameters to give a broad sense of how each of them affects the optimal portfolio.

Our base case is defined by the following parameters:

- Required real return: 4%;
- Investment horizon length: 60 months; and
- Expected block length: 36 months (i.e. the "intermediate" block length discussed in Section 4.3.4).

5.1.2. The optimal portfolios for all problem settings (employing intermediate block lengths)

Our final model here is specified with reference to the chosen "intermediate" expected block lengths discussed in Section 4.3.4.

The optimal portfolios are derived assuming those block lengths represent a compromise between bias and variance. Table 18 and Table 19 below show these optimal portfolios for the five asset case:

criterion	horizon	req real return	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	long run return	annualised monthly stdev
MVO (req ret)		2%	2%	3%	90%	1%	4%	3%	5%	2.0%	1.6%
MVO (req ret)		4%	21%	9%	49%	1%	19%	23%	20%	4.0%	5.4%
MVO (req ret)		6%	41%	7%	10%	7%	35%	48%	42%	6.0%	10.4%
prob of success	12	2%	49%	0%	7%	18%	27%	66%	45%	6.7%	12.6%
prob of success	36	2%	54%	9%	4%	0%	34%	54%	34%	6.5%	12.1%
prob of success	60	2%	55%	13%	0%	0%	32%	55%	32%	6.5%	12.3%
prob of success	84	2%	70%	2%	2%	0%	26%	70%	26%	7.2%	15.0%
prob of success	120	2%	64%	0%	0%	1%	35%	65%	36%	7.1%	14.1%
prob of success	240	2%	64%	0%	1%	0%	35%	64%	35%	7.1%	14.0%
prob of success	12	4%	53%	0%	0%	38%	9%	91%	47%	7.2%	15.1%
prob of success	36	4%	68%	0%	1%	16%	16%	84%	32%	7.4%	15.7%
prob of success	60	4%	66%	2%	0%	0%	31%	66%	31%	7.1%	14.3%
prob of success	84	4%	73%	0%	0%	0%	26%	74%	26%	7.3%	15.6%
prob of success	120	4%	72%	0%	0%	2%	26%	74%	28%	7.3%	15.4%
prob of success	240	4%	72%	0%	0%	0%	28%	72%	28%	7.3%	15.4%
prob of success	12	6%	68%	0%	0%	32%	0%	100%	32%	7.6%	17.0%
prob of success	36	6%	56%	0%	0%	43%	1%	99%	44%	7.4%	16.0%
prob of success	60	6%	77%	0%	0%	6%	16%	84%	22%	7.5%	16.8%
prob of success	84	6%	71%	0%	0%	21%	8%	92%	29%	7.6%	16.7%
prob of success	120	6%	85%	0%	0%	15%	0%	100%	15%	7.7%	18.9%
prob of success	240	6%	79%	0%	0%	14%	7%	93%	21%	7.6%	17.7%
avg shortfall	12	2%	11%	0%	84%	0%	5%	11%	5%	2.7%	2.7%
avg shortfall	36	2%	26%	0%	60%	0%	14%	26%	14%	4.1%	5.8%
avg shortfall	60	2%	34%	0%	47%	0%	19%	34%	19%	4.8%	7.6%
avg shortfall	84	2%	49%	0%	29%	0%	21%	49%	21%	5.8%	10.6%
avg shortfall	120	2%	52%	0%	21%	0%	26%	52%	26%	6.2%	11.4%
avg shortfall	240	2%	59%	0%	8%	0%	33%	59%	33%	6.7%	12.9%
avg shortfall	12	4%	24%	4%	57%	0%	15%	24%	15%	4.0%	5.5%
avg shortfall	36	4%	35%	0%	46%	0%	19%	35%	19%	4.8%	7.7%
avg shortfall	60	4%	47%	0%	26%	0%	27%	47%	27%	5.8%	10.4%
avg shortfall	84	4%	61%	0%	7%	0%	31%	61%	31%	6.8%	13.3%
avg shortfall	120	4%	66%	0%	0%	0%	34%	66%	34%	7.1%	14.3%
avg shortfall	240	4%	68%	0%	0%	0%	32%	68%	32%	7.2%	14.7%
avg shortfall	12	6%	42%	14%	2%	0%	43%	42%	43%	6.0%	10.5%
avg shortfall	36	6%	48%	2%	18%	0%	32%	48%	32%	6.0%	10.7%
avg shortfall	60	6%	60%	3%	0%	0%	37%	60%	37%	6.9%	13.3%
avg shortfall	84	6%	68%	0%	0%	0%	32%	68%	32%	7.2%	14.6%
avg shortfall	120	6%	70%	0%	0%	0%	30%	70%	30%	7.3%	15.0%
avg shortfall	240	6%	73%	0%	0%	1%	26%	74%	27%	7.4%	15.7%
ds dev	12	2%	8%	0%	89%	2%	1%	10%	3%	2.4%	2.1%
ds dev	36	2%	21%	0%	67%	0%	12%	21%	12%	3.6%	4.7%
ds dev	60	2%	27%	0%	58%	0%	15%	27%	15%	4.2%	6.1%
ds dev	84	2%	39%	0%	44%	0%	17%	39%	17%	5.0%	8.4%
ds dev	120	2%	46%	0%	32%	0%	23%	46%	23%	5.6%	9.9%
ds dev	240	2%	54%	0%	16%	0%	29%	54%	30%	6.3%	11.9%
ds dev	12	4%	22%	1%	57%	2%	17%	25%	19%	4.0%	5.5%
ds dev	36	4%	25%	0%	60%	0%	15%	25%	15%	4.0%	5.6%
ds dev	60	4%	35%	0%	47%	0%	19%	35%	19%	4.8%	7.6%
ds dev	84	4%	49%	0%	27%	0%	24%	49%	24%	5.9%	10.5%
ds dev	120	4%	58%	0%	12%	0%	30%	58%	30%	6.6%	12.6%
ds dev	240	4%	65%	0%	0%	0%	35%	65%	35%	7.1%	14.1%
ds dev	12	6%	43%	6%	11%	2%	38%	45%	40%	6.0%	10.5%
ds dev	36	6%	43%	0%	14%	0%	42%	43%	42%	6.0%	10.5%
ds dev	60	6%	47%	0%	21%	0%	32%	47%	32%	6.0%	10.6%
ds dev	84	6%	58%	0%	11%	0%	31%	58%	31%	6.6%	12.6%
ds dev	120	6%	65%	0%	0%	0%	35%	65%	35%	7.1%	14.2%
ds dev	240	6%	70%	0%	0%	0%	30%	70%	30%	7.3%	15.0%

Table 18: All optimal portfolios for intermediate block lengths for all problem settings

criterion	horizon	req real return	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	long run return	annualised monthly stdev
max drawdown	12	2%	1%	4%	89%	4%	2%	5%	6%	2.0%	1.6%
max drawdown	36	2%	1%	3%	90%	4%	2%	5%	6%	2.0%	1.6%
max drawdown	60	2%	2%	0%	92%	0%	6%	2%	6%	2.0%	1.6%
max drawdown	84	2%	2%	0%	92%	0%	6%	2%	6%	2.0%	1.6%
max drawdown	120	2%	3%	1%	92%	0%	5%	3%	5%	2.0%	1.6%
max drawdown	240	2%	2%	0%	91%	0%	6%	2%	6%	2.0%	1.7%
max drawdown	12	4%	21%	1%	55%	2%	22%	23%	23%	4.0%	5.5%
max drawdown	36	4%	20%	4%	50%	0%	25%	21%	25%	4.0%	5.6%
max drawdown	60	4%	21%	10%	45%	0%	25%	21%	25%	4.0%	5.6%
max drawdown	84	4%	20%	11%	44%	0%	25%	20%	25%	4.0%	5.6%
max drawdown	120	4%	20%	5%	52%	4%	20%	24%	24%	4.0%	5.5%
max drawdown	240	4%	19%	20%	34%	0%	26%	19%	27%	4.0%	5.7%
max drawdown	12	6%	41%	11%	5%	1%	42%	43%	43%	6.0%	10.5%
max drawdown	36	6%	40%	4%	6%	0%	50%	40%	50%	6.0%	10.7%
max drawdown	60	6%	40%	10%	2%	0%	47%	41%	47%	6.0%	10.6%
max drawdown	84	6%	44%	4%	13%	0%	38%	44%	38%	6.0%	10.5%
max drawdown	120	6%	41%	7%	4%	0%	48%	41%	48%	6.0%	10.7%
max drawdown	240	6%	42%	3%	10%	0%	46%	42%	46%	6.0%	10.6%
sd horison	12	2%	2%	4%	90%	3%	0%	6%	3%	2.0%	1.6%
sd horison	36	2%	9%	9%	83%	0%	0%	9%	0%	2.4%	2.4%
sd horison	60	2%	10%	13%	77%	0%	0%	10%	0%	2.5%	2.8%
sd horison	84	2%	13%	16%	70%	0%	0%	13%	0%	2.8%	3.5%
sd horison	120	2%	14%	18%	68%	0%	0%	14%	0%	2.9%	3.8%
sd horison	240	2%	18%	25%	57%	0%	0%	18%	0%	3.2%	4.8%
sd horison	12	4%	23%	10%	51%	0%	16%	23%	16%	4.0%	5.5%
sd horison	36	4%	24%	24%	41%	0%	11%	24%	11%	4.0%	6.0%
sd horison	60	4%	25%	31%	36%	0%	8%	25%	8%	4.0%	6.4%
sd horison	84	4%	27%	35%	35%	0%	3%	27%	3%	4.0%	6.9%
sd horison	120	4%	29%	25%	46%	0%	0%	29%	0%	4.0%	6.9%
sd horison	240	4%	29%	28%	43%	0%	0%	29%	0%	4.0%	7.0%
sd horison	12	6%	44%	3%	15%	1%	37%	45%	38%	6.0%	10.5%
sd horison	36	6%	47%	16%	7%	0%	30%	47%	30%	6.0%	10.8%
sd horison	60	6%	47%	25%	0%	0%	28%	47%	28%	6.0%	10.9%
sd horison	84	6%	50%	24%	4%	0%	22%	50%	22%	6.0%	11.3%
sd horison	120	6%	51%	29%	1%	0%	19%	51%	19%	6.0%	11.6%
sd horison	240	6%	58%	9%	26%	0%	7%	58%	7%	6.0%	12.4%

Table 19: All optimal portfolios for intermediate block lengths for all problem settings (continued)

It is important to realise that each and every one of the portfolios in Table 18 and Table 19 is optimal, depending on the exact problem setting. These optimal portfolios will be discussed in the following sections, but the following is immediately evident from Table 18 and Table 19:

- the bespoke optimal portfolios are generally increasingly aggressive in the investment horizon (the probability of success portfolio is an exception);
- the maximum drawdown portfolio is quite similar to the MVO portfolio; and
- the following order of aggressiveness (in terms of both total equity weight and monthly standard deviation) can generally be observed in the various optimal portfolios:

  probability of success > average shortfall > downside deviation > sd horizon > maximum drawdown portfolio > MVO

5.1.3. Discussion of the optimal portfolios

The graph below shows the optimal portfolios belonging to all the risk criteria for our base case problem setting. The various risk criteria result in markedly different optimal portfolios in the same problem setting. Figure 27 below shows the optimal weights for the various criteria for our base case setting:

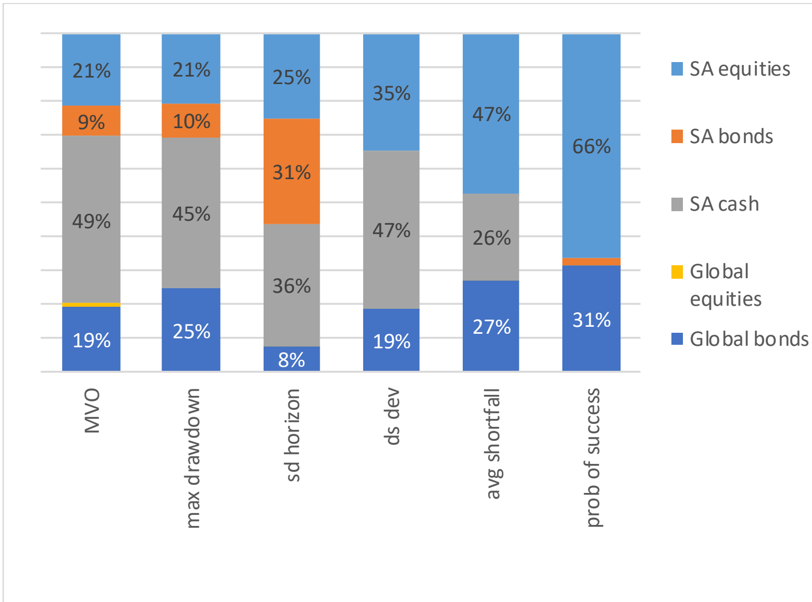

Figure 27: Optimal weights for the base case (i.e. horizon = 60 months, required real return = 4%)

As one may expect from the inspection of the weights of the optimal equity portfolios above, the portfolios of Figure 27 occupy very different positions on the mean-variance continuum, as is depicted in Figure 28:

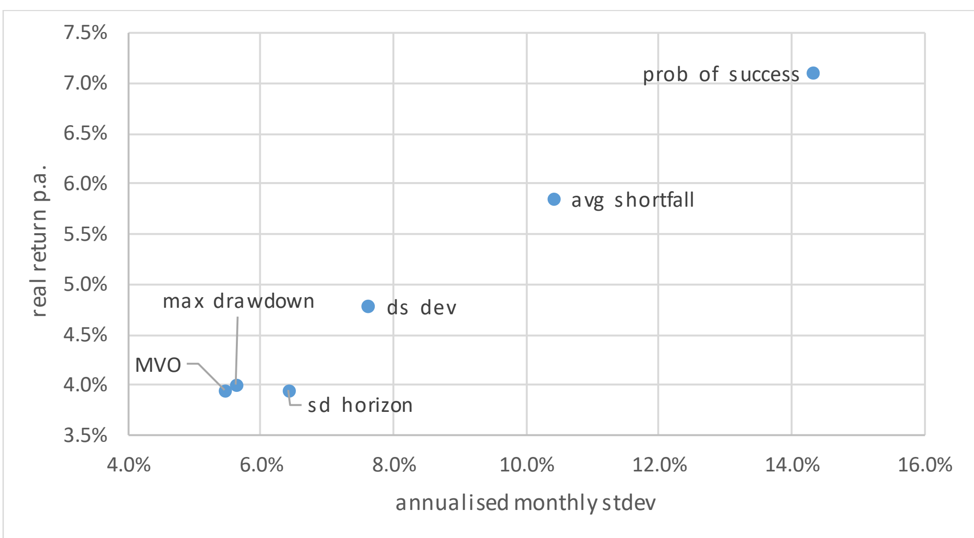

Figure 28: Real return versus annualised monthly standard deviation for the base case (horizon=60 months, required real return = 4%)

Not surprisingly, considering the penalty function implied by these criteria, Figure 27 and Figure 28 show that the aggressiveness of the optimal portfolios as measured by the total equity allocation and the monthly standard deviation (for a given level of required real return and investment horizon) can generally be ordered as follows:

> probability of success > average shortfall > downside deviation > sd horizon > maximum drawdown portfolio > MVO

Figure 27 and Figure 28 thus make clear that the choice of risk parameter has material consequences for the optimal portfolio. An investor (with a required real return of 4% and investment horizon of 60 months and a utility function consistent with those underlying either the average shortfall or downside deviation risk criteria) would be severely underexposed to growth assets if he had invested in the MVO portfolio with average return equal to the required real return. The portfolio with the highest probability of success in achieving a real return of 4% over 60 months is even more aggressive than the average shortfall and downside deviation portfolios.

Comparison to MVO

It is important to point out that these portfolios are not merely different points on the mean-variance efficient frontier: apart from the MVO solution, none of these portfolios are on the traditional efficient frontier based on the typical monthly periodicity. Table 20 shows the weight of each optimal portfolio above minus the weight of the MVO portfolio with the same average real return:

criterion	avg real return	annualised monthly stdev	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore
prob of success	7.1%	14.3%	9%	2%	0%	-16%	5%	-7%	-12%
avg shortfall	5.8%	10.4%	7%	-3%	6%	-9%	-2%	-2%	-10%
ds dev	4.8%	7.6%	5%	-2%	5%	-6%	-2%	-1%	-8%
sd horizon	4.0%	6.4%	4%	30%	-22%	-5%	-7%	-1%	-12%
max drawdown	4.0%	5.6%	-1%	9%	-13%	-5%	10%	-6%	5%

*Table 20: The optimal portfolios of the **base case** versus MVO portfolio with the same average real return. The figures shown under the asset class are the weights of the optimal portfolios minus the weights of the MVO optimal portfolios.*

It is perhaps more amenable to interpretation to consider the averages (within risk criterion) of these differences across all the possible problem settings defined in 5.1.1, depicted in Table 21:

	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	% improvement vs. MVO (avg return)
average for all prob of success	5%	1%	0%	-8%	2%	-3%	-6%	1%
average for all avg shortfall	7%	0%	2%	-11%	2%	-4%	-9%	11%
average for all ds dev	6%	-1%	3%	-9%	1%	-3%	-7%	7%
average for all sd horizon	5%	18%	-11%	-5%	-6%	0%	-12%	9%
average for all drawdowns	2%	4%	-4%	-3%	11%	-1%	8%	9%

*Table 21: The optimal portfolios of **all problem settings** versus MVO portfolio with the same average real return. The figures shown under the asset class are the optimal portfolios minus the MVO optimal portfolios.*

The most notable and consistent patterns in Table 21 (the standard deviation over the horizon portfolio being an exception) are that local equities and global bonds receive higher allocations, at the expense of global equities, and that the total offshore allocation reduces overall.

A higher SA equities allocation at the expense of global equities makes sense from an economic perspective: South Africa's equity returns should more consistently be related to SA inflation, as the SA economy underlies both (but is only a very small component of the global economy). This relationship may not be manifested in month-on-month fluctuations, but arguably is manifested over the longer term and is captured by longer block lengths. This is consistent with the long-term mean reversion displayed by SA equities returns causing a monotone (in periodicity) decreasing ratio of SA equities returns standard deviation versus that of global equities in ZAR terms (discussed in Section 4.1.4.2). It is also consistent with the purported phenomenon that investors tend to invest more in local equities than offshore equities (see for example Ang & Bekaert, 2002).

The reduction of global equities is in itself perhaps the reason for an increased global bonds exposure: To regain some the diversification benefits of an exposure to offshore assets and exchange rates, the allocation to global bonds increases. As discussed in Section 4.1.4.2, local equities, global bonds and global equities become increasingly attractive in terms of downside deviation and average downside against SA CPI as the periodicity increases. However, SA equities and global equities are highly correlated, and as already mentioned, SA equities dominate all asset classes as the periodicity increases. This perhaps explains why global equities lose some of its weight to local equities. Global bonds, on the other hand, are a far better diversifier than global equities, and hence it generally maintains its allocation.

Another striking result in Table 21 above is exhibited by only the standard deviation over investment horizon optimal portfolio: SA cash is down-weighted considerably and most of that allocation is gained by SA bonds. As discussed in Section 4.1.4.2 and illustrated by Figure 18 and Figure 19 on page 57, the advantage of the lower short-term variability of cash is greatly diminished over longer periodicities: bonds and cash deliver more and more similar returns the longer the period, but importantly, bonds outperformed on average, and hence start to dominate cash from a risk-adjusted return perspective. More generally the standard deviation over horizon comparison to MVO is an estimate of the bias in the weights of more traditional MVO portfolios due to *only* the time-dependencies in asset returns.

Interestingly, in the maximum drawdown portfolio one may have expected higher allocation to SA cash (the only asset class that has historically experienced no drawdowns over any length of time in rand terms) at the expense of perhaps SA bonds, but the opposite is the case.

However, the question arises whether veering from the mean-variance efficient frontier *materially* improves the optimality of the portfolios. The final column in Table 21 above shows, in percentage terms, how much the bespoke optimal portfolios improve on average across problem settings upon the MVO portfolio with the same average return. The improvements are material but modest. By contrast the improvements made relative to the MVO portfolio that merely satisfies the required real return, are quite considerable on average for all the criteria where the average return of the optimal portfolio tends to be higher than the required return. This is made clear by Table 22, which compares the average percentage improvement relative to these two MVO portfolios:

	vs MVO (same return)	vs MVO (req return)
prob of success	1%	44%
avg shortfall	11%	34%
ds dev	7%	22%
sd horizon	9%	6%
max drawdown	9%	26%

Table 22: Average improvement in the chosen risk criterion versus the MVO portfolios with respectively the same average return and with the required real return

Investment horizon

To emphasise the effect of investment horizon on the optimal portfolio, Figure 29 below plots the optimal total equity weight against this variable for five different risk criteria, for a required real return of 4%. In the case of average shortfall, downside deviation, and to a lesser extent standard deviation over the horizon, the optimal equity allocation, real return and monthly standard deviation are all clearly increasing in the length of the investment horizon:

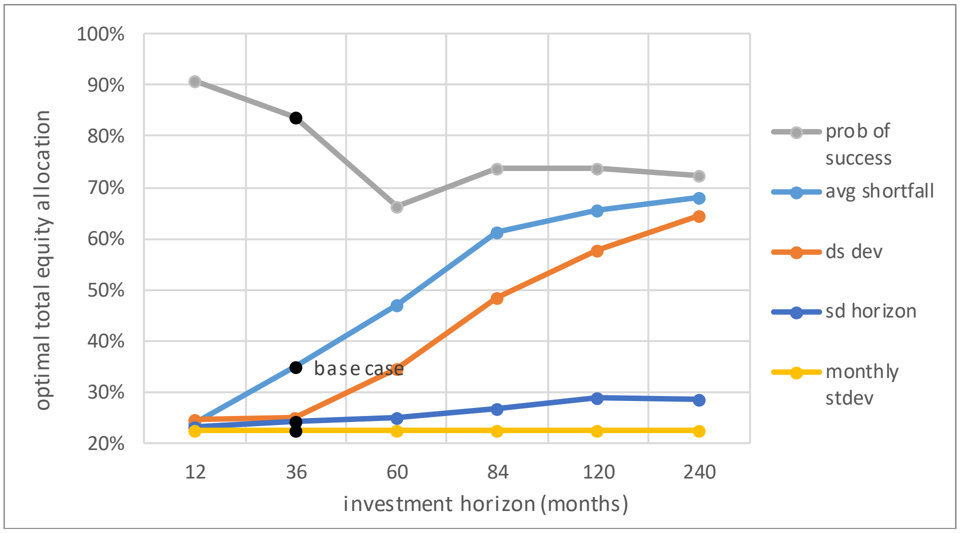

Figure 29: The total equity allocation of the optimal portfolios (belonging to all of the risk criteria) versus the investment horizon length (required real return of 4%)

Figure 30 and Figure 31 plot the same information as Figure 29, but the vertical axes depict respectively monthly standard deviation and real return instead of the optimal total equity allocation:

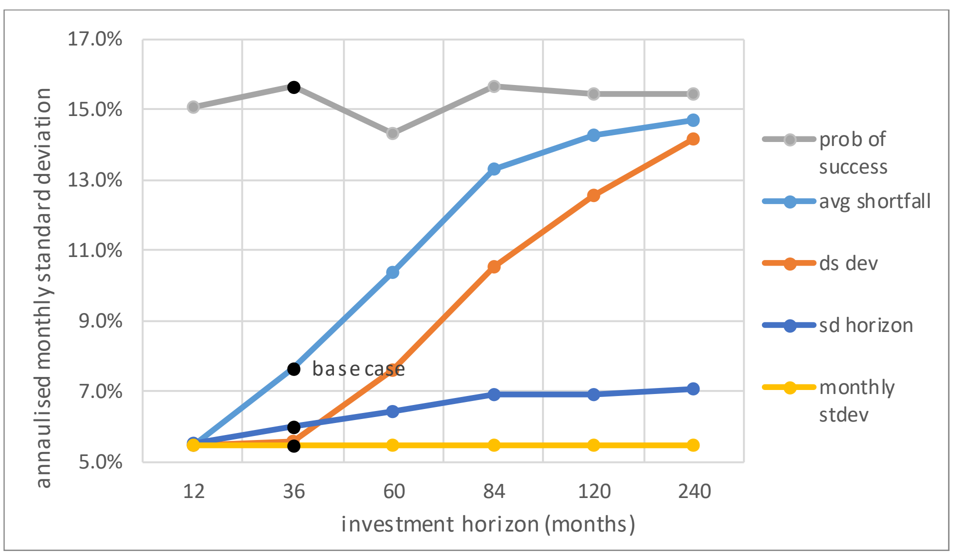

Figure 30: The annualised monthly standard deviation of the optimal portfolios (belonging to all of the risk criteria) versus the investment horizon length (required real return of 4%)

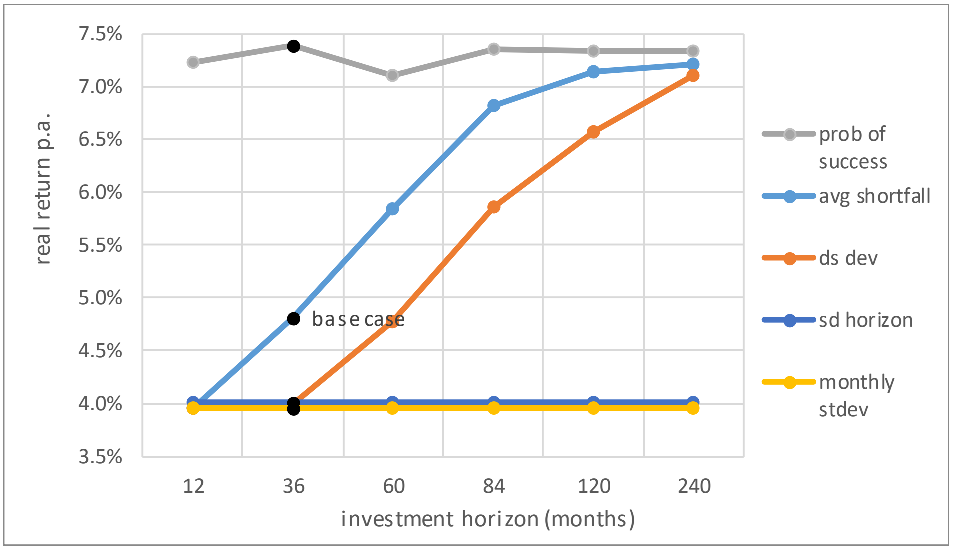

Figure 31: The real return p.a. of the optimal portfolios (belonging to all of the risk criteria) versus the investment horizon length (required real return of 4%)

Figure 29 and Figure 30 above reveal that, perhaps surprisingly, the optimal probability of success portfolio does not become more aggressive as the investment horizon increases (though not shown here, the actual probability of success does however increase from 62% to 93%).

The pattern above of increasing aggressiveness with increased investment horizon is consistent across real return targets. Consider for example the optimal total equity allocation for the average shortfall optimal portfolio for a required real return of 2%, 4% and 6% respectively, depicted in Figure 32:

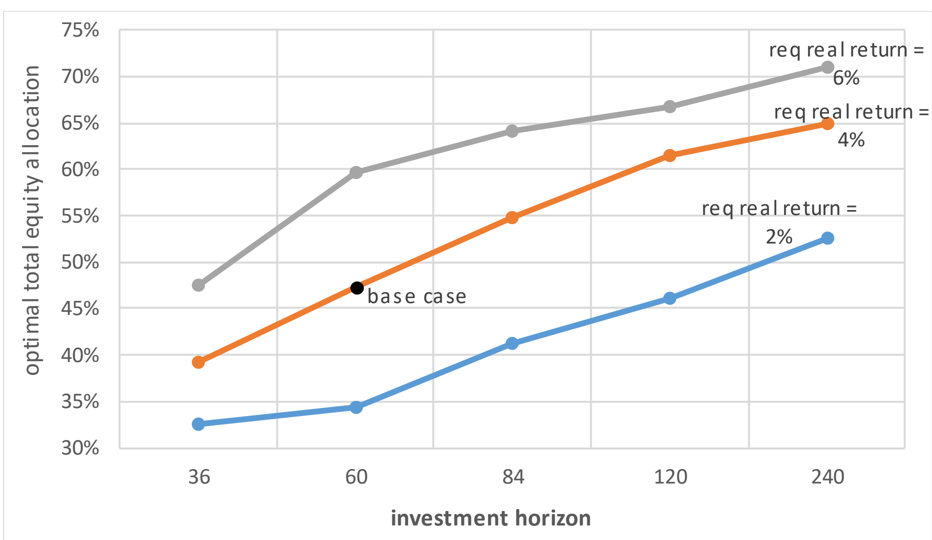

*Figure 32: The total equity allocation of the optimal portfolios (for the **average shortfall portfolio**) versus the investment horizon length*

As already touched on earlier, the optimal weight to SA bonds is generally increasing in the investment horizon for the standard deviation over horizon portfolio, as can be seen in Figure 33, which plots the optimal SA bonds weight against investment horizon for all three required real returns for this criterion:

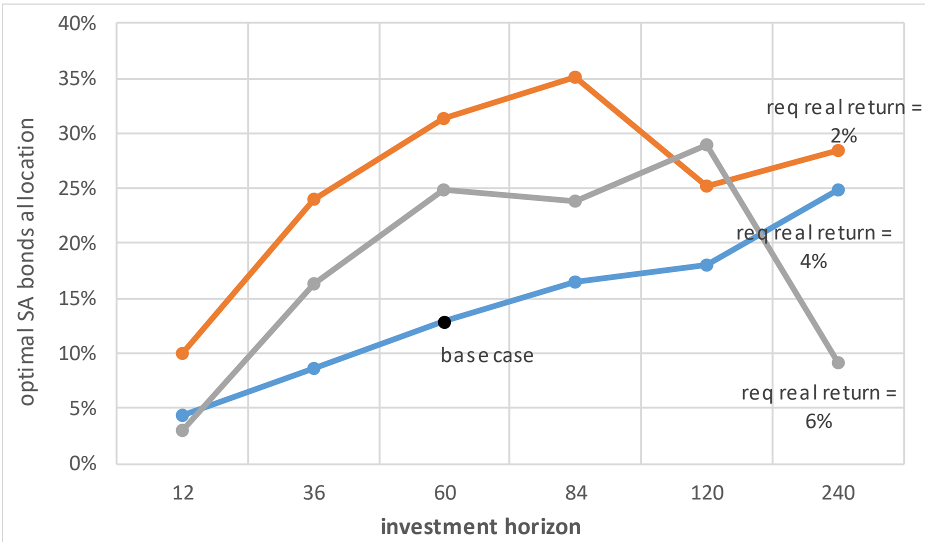

Figure 33: The optimal SA bonds allocation of the standard deviation over horizon portfolio versus investment horizon length

The probabilities of achieving the targets are also interesting in and of themselves. Figure 34 below plots the probability of success of the optimal probability of success portfolios against investment horizon, for each of the required real return targets. We can see how the probability of achieving the required real return increases with the investment horizon and decreases with the required real return. A comparison of the solid lines to the dashed lines is an estimate of the increase in probability of achieving the target due to the time-dependencies in asset returns. It should be borne in mind that the MVO portfolio that merely targets the required real return will have, roughly speaking, only a 50% probability of achieving the target over any investment horizon.

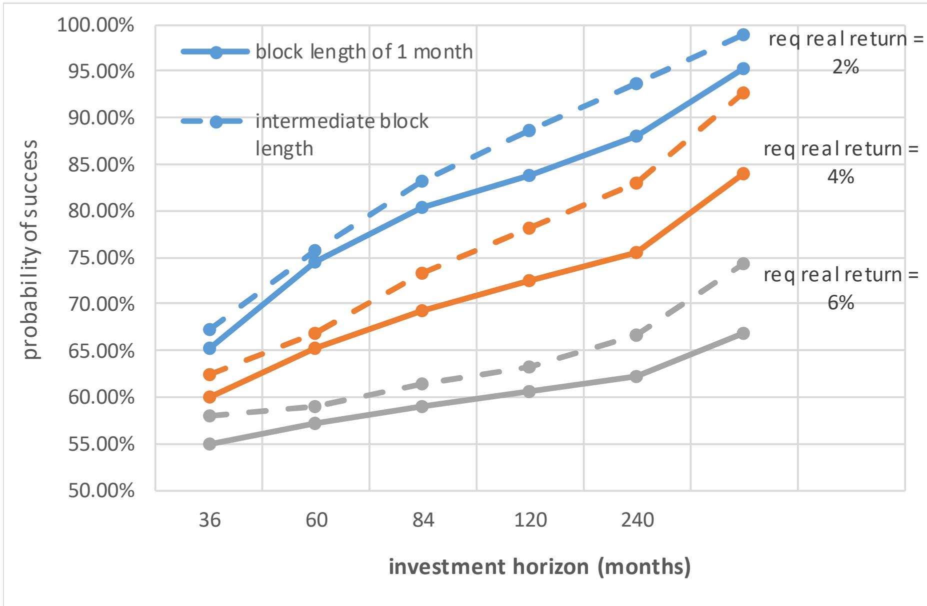

Figure 34: A plot of the probability of success against the investment horizon length. The probability increases quickly with investment horizon. The difference between the solid line (1-month block length) and dashed line (intermediate block length) is an estimate for the increase in probability due mean reversion in local equity and other time-dependencies in asset returns captured by the longer block length.

By comparing the optimal allocations of the average shortfall risk criterion for a block length of one month to those of the "intermediate" block lengths, we can isolate the effect of time-dependencies in returns. The balance of the increase in equities (relative to the MVO portfolio, which is invariant to investment horizon) could then be considered as related to the lack of an upside variance penalty. This allows us to split the increase in equity (versus MVO) into the portion attributable to the time-dependencies in returns (arguably chiefly long-term mean reversion in SA equities), and the portion attributable to the lack of upside variance penalty inherent to the average shortfall portfolio. This is depicted in Figure 35:

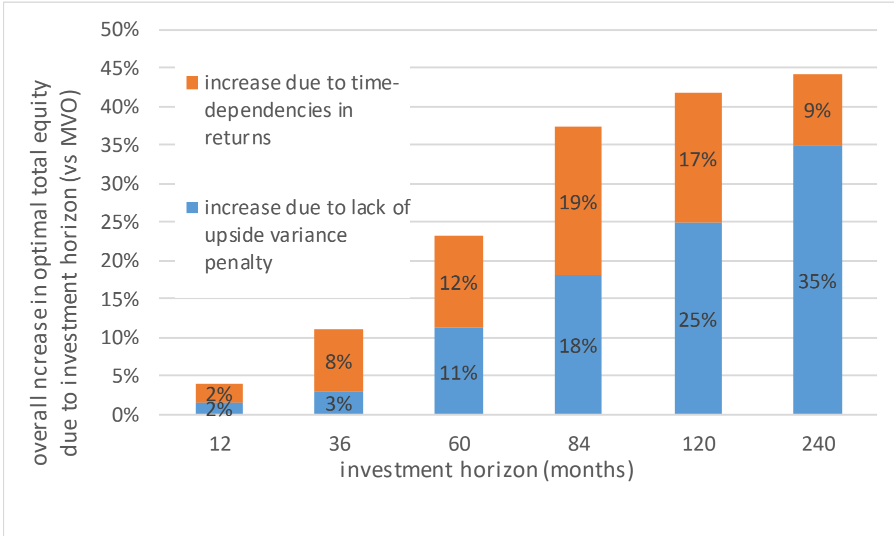

*Figure 35: Increase in total equity allocation versus investment horizon due to time dependencies and lack of penalty on upside variation for an investor averse **to average shortfall** and **required real return of 4%***

For investors who are averse to both upside and downside variation, a more relevant estimate of the increase in the optimal allocation due to time-dependencies in asset returns can be isolated by comparing the standard deviation over horizon portfolio against the MVO portfolio. This estimate is depicted for the various investment horizons in Figure 36

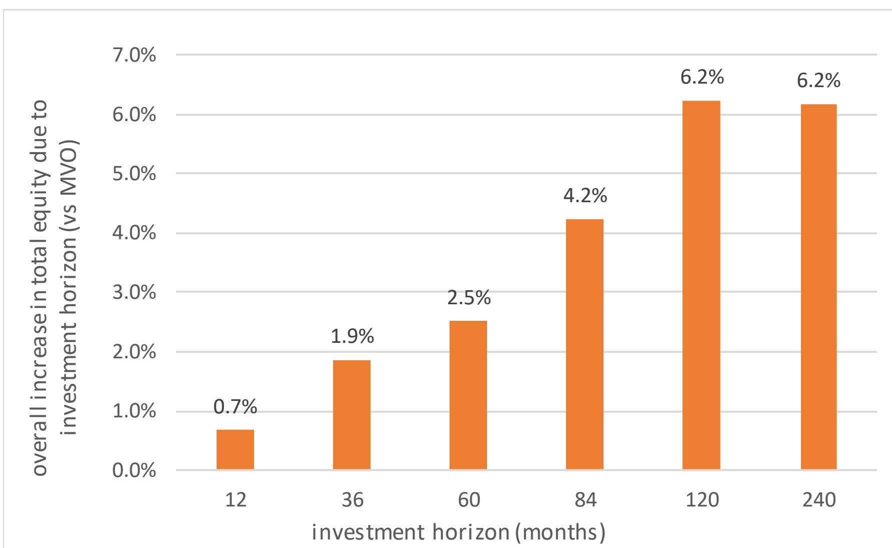

*Figure 36: Increase in total equity allocation due to time-dependencies only (from the point of view of an investor averse to both upside and downside variation) versus investment horizon (**required real return of 4%)***

Comparing Figure 35 and Figure 36, one can see that the increase in optimal equity weight owing to time-dependencies of returns is much smaller when upside and downside variations around the mean is penalised (as opposed to only downside variance relative to CPI+4%), but it is still material.

The increase in total equities attributed to the time-dependencies in returns is, in both cases above, likely mainly due to the long-term mean reversion of SA equities returns described in Section 4.1.4.2. Consistent with the fact that mean reversion is exhibited by SA equities (but not in global equities in ZAR terms), the increasing optimal total equity weight due to time-dependencies as the investment horizon increases can be observed to solely due to an increase in the weight of SA equities, rather than global equities.

Finally, we show in Figure 37 the total equity allocations of the optimal portfolios for standard deviation over horizon, downside deviation and average shortfall for all three required real returns and for all investment horizons:

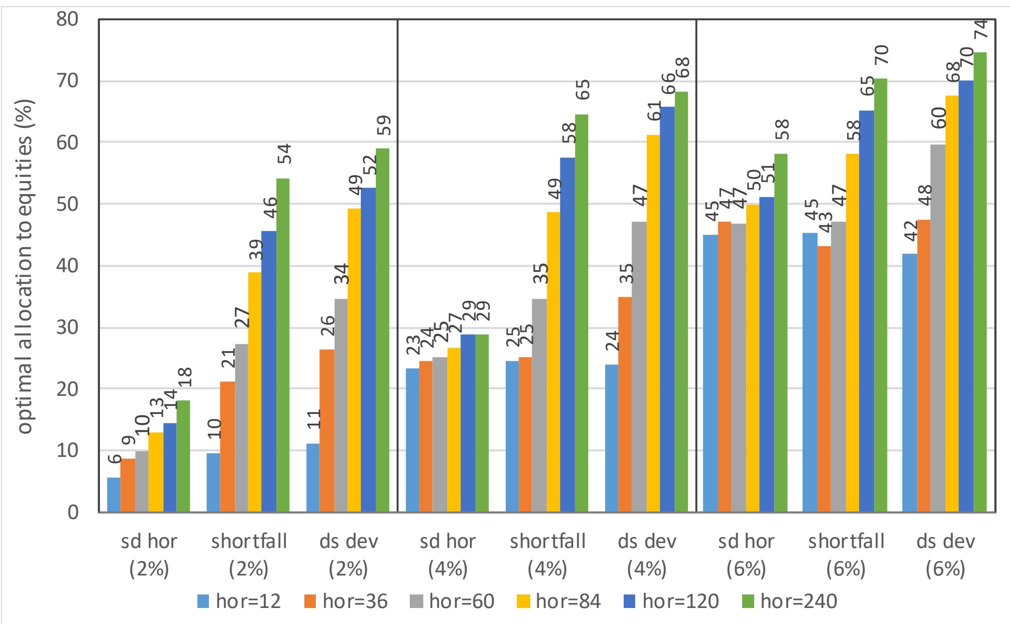

Figure 37: The optimal total equity allocation for four of the risk criteria, the three different required real returns (2%, 4% and 6%) and for all investment horizons. Though we have omitted the MVO portfolio, the standard deviation over horizon portfolio for an investment horizon of 12 months is a reasonable proxy.

A quick inspection of Figure 37 shows that the effect of required real return, risk criterion, and investment horizon all materially impact the total equity allocation with broadly comparable orders of magnitude.

Block length

In Section 4.3.4 we discussed the issue of the block length. It is a nuisance parameter of our model that is difficult to determine, hence it is important to discuss the effect of its choice on optimal portfolios.

The optimal portfolios discussed so far have focused on the intermediate block lengths, which represent a compromise between bias and variance. Two alternative block lengths are of particular importance and significance: a block length of one month and a block length equal to the investment horizon length (the latter will be referred to as the *empirical case*).

A block length of one month is important since, as discussed in Section 4.3.4, it is the block length that minimises estimates of the standard error of the variance of returns over the investment horizon (although with a gain in bias). It is also the block length that treats monthly returns as i.i.d., as is typically the case in practice.

The empirical block length on the other hand is important as the resultant optimal portfolios document which portfolios *actually* performed the best historically for the period 1972 to 2017. Though we will only touch on the results of a one month block length in this section, the optimal portfolios belonging to both of these block lengths are tabled in full in Appendices 7.2.1 and 7.2.2.

Table 23 below shows the optimal allocations for all the risk criteria for a block length of one month minus the optimal portfolios for the intermediate block length:

criterion	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	avg real return	annualied monthly stdev
avg shortfall	-18%	5%	11%	4%	-3%	-13%	1%	-1.0%	-3%
ds dev	-13%	3%	9%	3%	-2%	-10%	1%	-0.8%	-2%
max drawdown	0%	-5%	6%	0%	0%	0%	0%	0.0%	0%
prob of success	-9%	-2%	0%	15%	-4%	6%	11%	0.0%	0%
sd horison	-5%	-11%	1%	0%	15%	-5%	15%	0.0%	-1%

Table 23: The change in optimal weights, real return and monthly standard deviation if the block length changes from "intermediate" block lengths to one month (required real return of 4% and 60 month investment horizon).

The generally lower allocation to SA equities for the one-month block length evident from Table 23, as well as the commensurate decrease in monthly standard deviation and real return, must be solely due to the time-dependencies of asset returns over periods longer than one month. The numbers in Table 23 can thus be considered estimates for the effect of these phenomena on optimal portfolios. Figure 38 below depicts the optimal SA equities allocation for the average shortfall criterion for various block lengths and a required real return of 4%. One can see how, for this criterion and required real return, the equity allocation gradually increases as the block length increases (capturing the long-term mean reversion in SA equities):

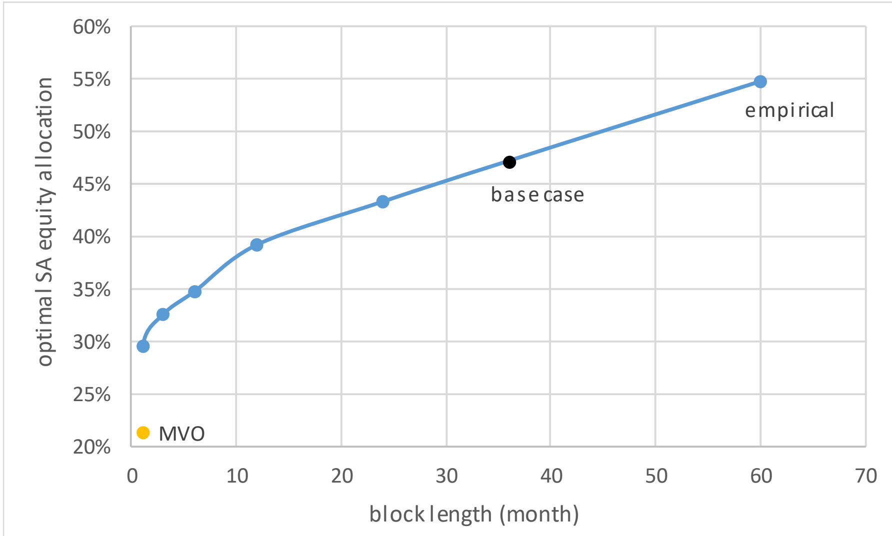

Figure 38: The optimal equity allocation (for the ***average shortfall criterion*** *with a* ***required real return of 4%)*** *versus the block length*

5.2. Results of regime-cognisant optimal portfolios

In this section we discuss and compare the regime-cognisant optimal portfolios for a single investment horizon of 36 months and an arbitrary required real return of 6%. We confine ourselves to a narrower problem setting to conserve model running time, and for the sake of brevity.

Firstly, in Section 5.2.1, the effect of the tactical pressure parameter defined in Section 4.3.3.3 is discussed. The parameter is varied to illustrate how a practitioner could fine-tune the aggressiveness of his or her tactical positioning relative to the strategic asset allocation (i.e. the optimal allocation ignoring regimes or unconditionally optimal asset allocation).

In Section 5.2.2 we settle on a single value for the tactical pressure parameter, and compare the optimal portfolio weights (and their associated out-of-sample performance) belonging to each of our regime-schemes derived at six month intervals from 1972 to 2017 for the monthly standard deviation criterion.

Note than in Sections 5.2.1 and 5.2.2 our focus is narrowed on the monthly standard deviation criterion only (and a single arbitrary expected block length of 12 months), as this is the most familiar risk criterion, and is arguably most intuitively easy to understand and compare across regimes and values for the tactical pressure parameter.

In Section 5.2.3 we broaden our focus to include the optimal portfolio weights derived with respect to three additional risk criteria and two additional block lengths (one month and 24 months). We assess and compare the out-of-sample performance of the three different block lengths and the various regimes-schemes for four different risk criteria:

monthly standard deviation, standard deviation over horizon, average shortfall and maximum drawdown.

It may be useful for the reader to review the discussion (4.3.3.2) of the three different regime-schemes defined in Section 4.3.3:

- Regime-scheme 1: The state remains fixed on the latest state. The probability of resampling for every successive block is always related to the current (latest) state and is static. This probability is based on an inverse relationship with the multivariate distance between the current state and any other state.

- Regime-scheme 2: At the start of each bootstrap iteration the prevailing state is set to the current state at the date of the analysis, but thereafter changes within each iteration to the state at the end of each randomly resampled block.

- Regime-scheme 3: At the start of each bootstrap iteration, the prevailing state is set to the latest state available at the date of the analysis. The first resampled block then determines which historical date the simulation jumps to, and from there the state changes as per chronological history taking into account the block length of that iteration. The state then resets to the date of the analysis at the beginning of the next iteration. In essence, state changes occur as they had in history after the first resampled block in each iteration.

5.2.1. Effect of tactical pressure factor on optimal portfolios

As already discussed in 4.3.3.3 the tactical pressure parameter allows for more aggressive tactical tilting by rendering the resampling probabilities more unequal and therefore deviating further away from the "strategic" equal probabilities. If the pressure is 0, a regime-ignorant portfolio is retrieved, which could be considered the SAA from which we are tilting away when we increase the tactical pressure. Since we are employing monthly standard deviation as our risk parameter here, the SAA is simply the traditional MVO portfolio based on an equal weighting of all available monthly returns.

Figure 39, Figure 40 and Figure 41 below show the optimal portfolios for various tactical pressure parameters, for respectively regime-scheme 1, regime-scheme 2 and regime-scheme 3, for the arbitrary setting of an investment horizon of 36 months, an expected block length of 12 months, and required real return of 6%.[20] It is evident from these three figures how the optimal portfolios increasingly tilt away from the MVO portfolio as the tactical pressure is increased:

[20] The "current state" (as defined in Section 4.3.3.1) is set to 31 March 2017, and the optimal portfolios are found with reference to the entire data set (returns and regime-identifying variables from 1972 to 2017).

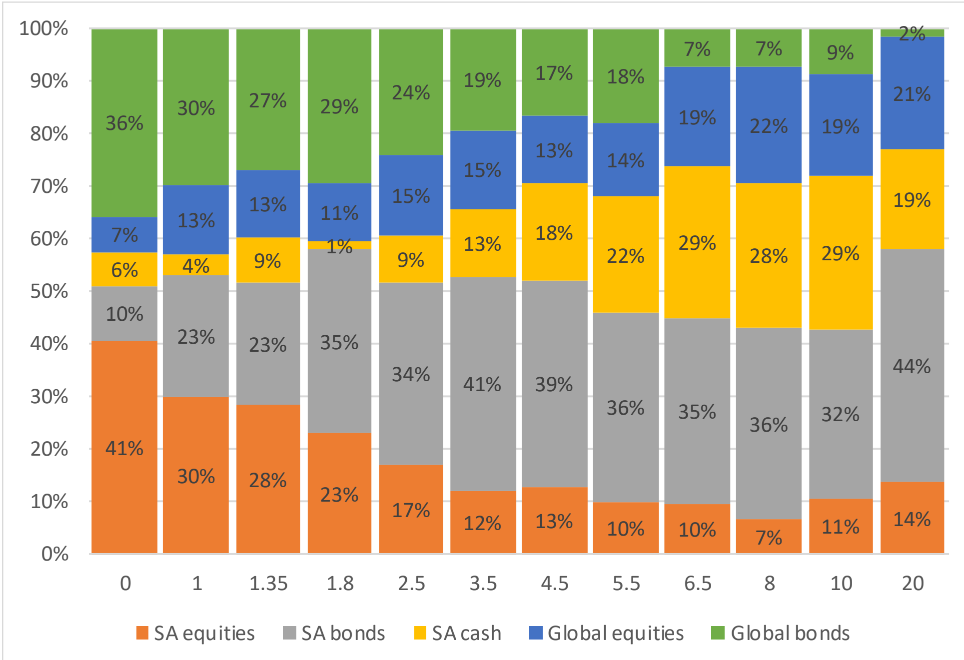

Figure 39: Optimal weights versus tactical pressure parameter for regime-scheme 1

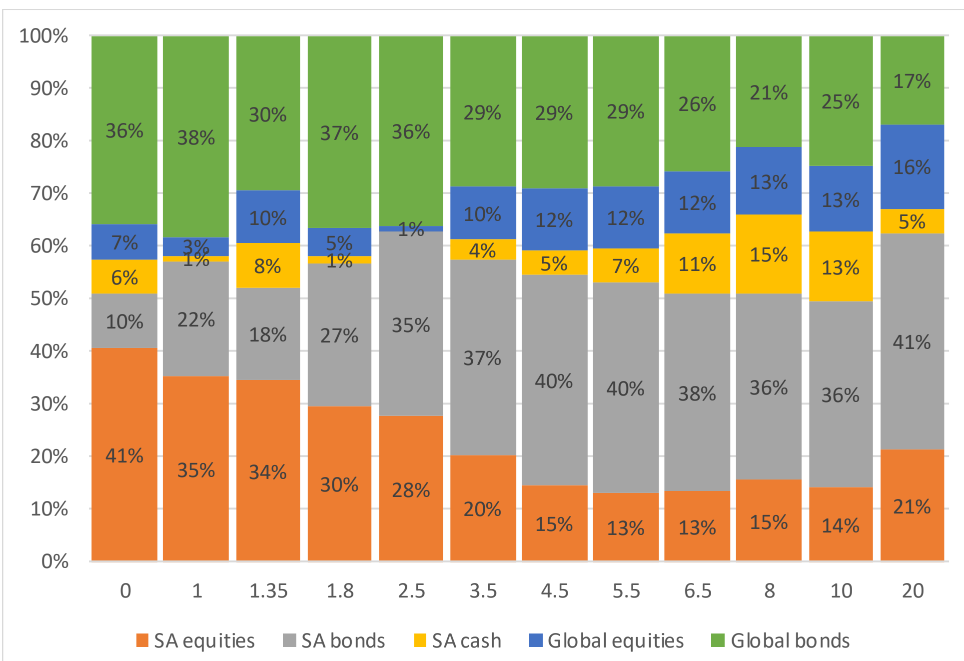

Figure 40: Optimal weights versus tactical pressure parameter for regime-scheme 2

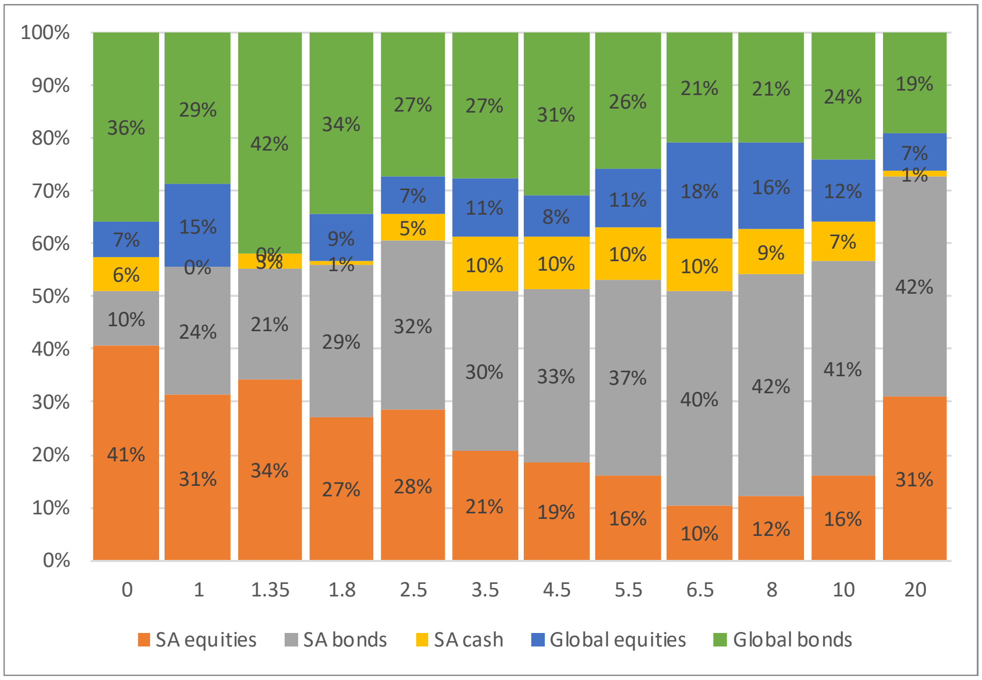

Figure 41: Optimal weights versus tactical pressure parameter for regime-scheme 3

5.2.2. Backtesting regime-cognisant portfolios over time (monthly standard deviation)

This section compares the optimal portfolios over time for the three different regime-schemes. For the sake of brevity, clarity and genetic algorithm running time considerations, it focuses on one very specific problem setting:

- Monthly standard deviation risk criterion only;
- Expected block length of 12 months;
- Horizon of 36 months;
- An aggressive tactical pressure parameter of 10;
- Required real return of 6% per annum; and
- Re-derive the optimal portfolio every six months using only data available at each point in time (which will enable us to perform out-of-sample performance backtesting).

As already discussed the monthly standard deviation risk criterion assumes that monthly returns are identically and independently distributed. If the variance of returns of a portfolio of assets is dependent on the periodicity (as the results in Section 4.1.4.2 suggest), monthly standard deviation is not a constant multiple of standard deviation over the investment horizon. However, we nevertheless confine ourselves to this risk measure in this particular section for the sake of simplicity, and to better explicate how the optimal portfolios that condition for our regime-schemes deviate from the typical

MVO approach. However, in Section 5.2.3 we will examine and compare the optimal portfolios belonging to some of the other risk measures, and backtest the performance for two additional block lengths.

Most of the discussion of the choice of block length in Section 4.3.4 still applies here. However, compared with the regime-ignorant case, the optimal block length can be argued to be shorter in this case: the unequal probabilities that are a function of the prevailing regime in themselves are hoped to account for the dependencies over consecutive months, and in some sense lessen the discontinuity between blocks (inspired by the matching block bootstrap).

The choice of a block length here is somewhat arbitrary but the fact that it is considerably smaller than the investment horizon is convenient. The fact that the investment horizon is three times as large as the expected block length means that the three different regime-schemes should give distinct optimal portfolios in each case. The state-changing mechanism would effectively be irrelevant if the expected block length is larger than the horizon, as the state would then rarely change. We could equally well optimise to any of the other criteria or maximise real return for a given level of risk, but instead focus on one risk parameter for the sake of brevity.

The tactical pressure parameter is deliberately very aggressive, to ensure that the three regime-changing mechanisms yield returns distribution that are materially dissimilar. Recall that the resultant weighting scheme for a tactical pressure of 10 has already been explicated by Table 12, Table 13, Figure 20 and Figure 21 (found between pages 78 to 88), and those items should be referred to for a clearer picture of the current setting.

Figure 42 to Figure 45 below are the optimal portfolios as at different points in time for respectively no regime-scheme, regime-scheme 1, regime-scheme 2 and regime-scheme 3. Importantly, as we are employing monthly standard deviation as our measure of risk, the regime-ignorant scheme is simply traditional MVO. The regime-ignorant scheme could also be considered the strategic asset allocation from which the three regime-cognisant portfolios tilt away – it is based at every point in time on all the data in our dataset that was available at the time. As a result it also changes over time as new data becomes available. The first optimal portfolio is thus based on the 11 years of data from March 1972 to March 1983. For every successive optimal portfolio, six months’ worth of additional data is included in the dataset (each optimal portfolio is based on more data than the previous optimal portfolio).

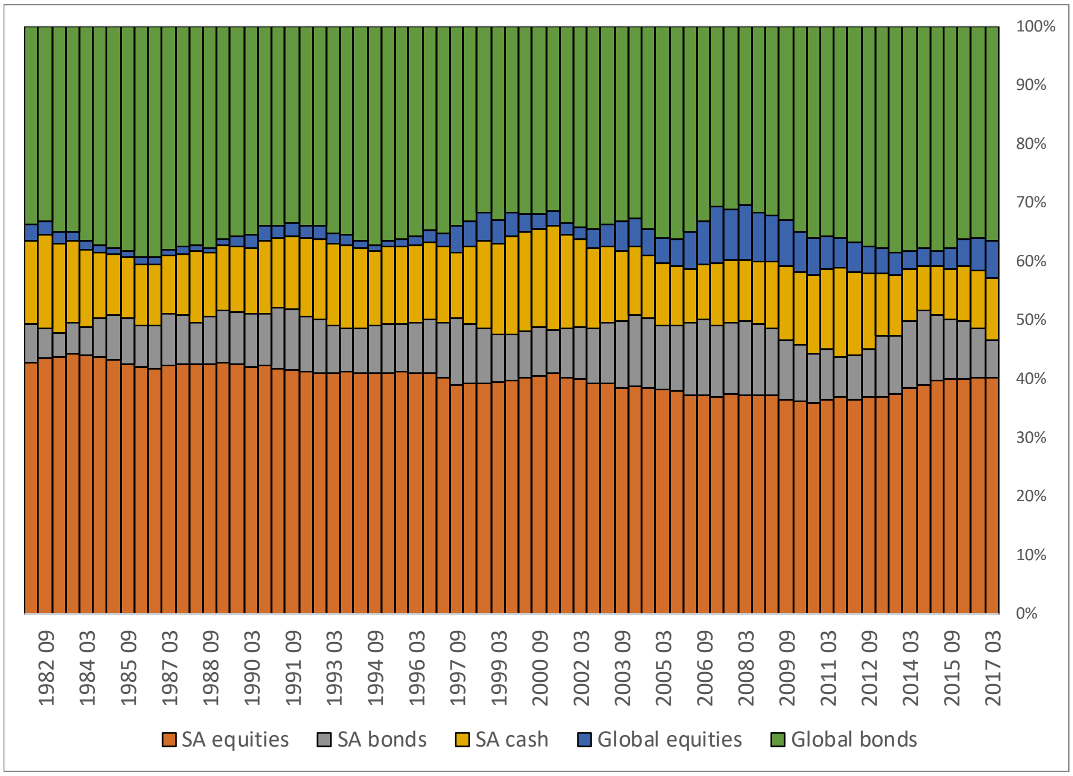

Figure 42: Optimal weights over time for no regime-scheme (i.e. MVO or the strategic asset allocation). The optimal allocation changes over time purely because new returns become available and are incorporated in the analysis.

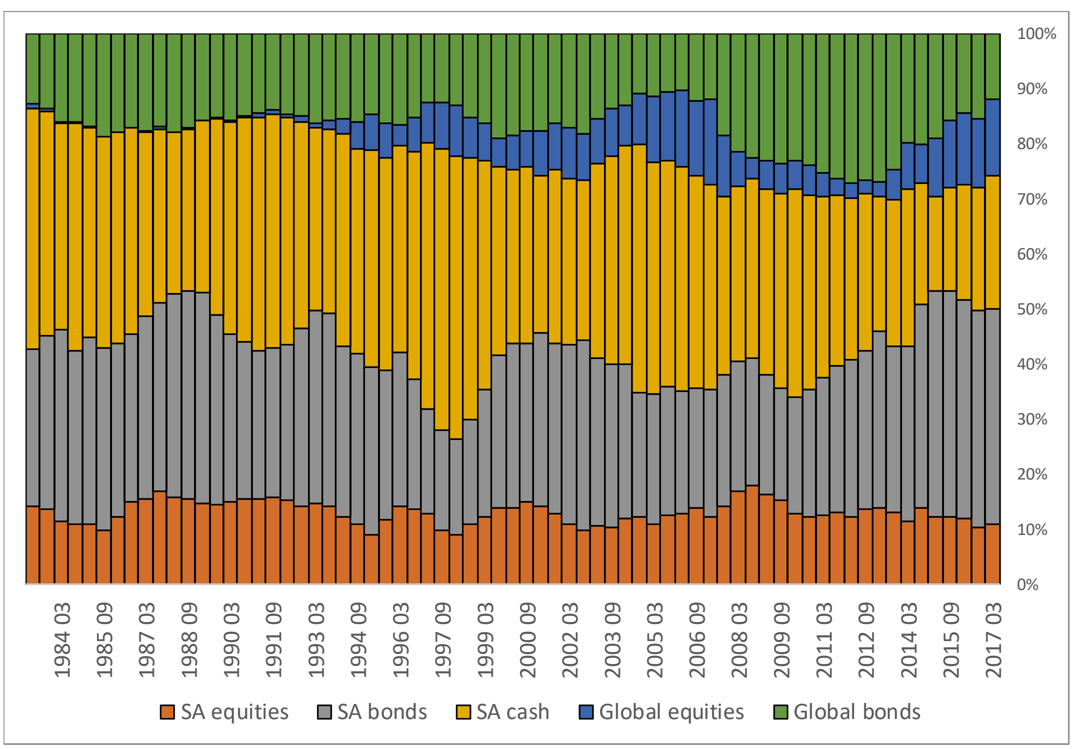

Figure 43: Optimal weights over time for regime-scheme 1. The optimal allocation changes as new returns data is incorporated and the regime-identifying variables are updated.

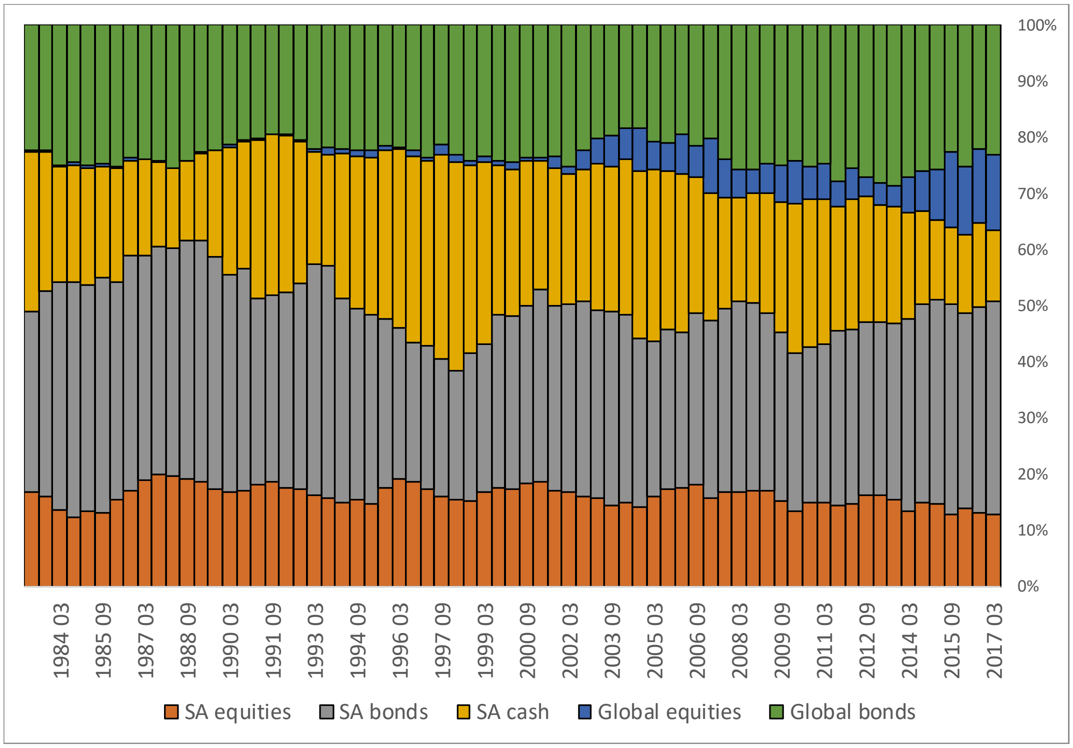

Figure 44: Optimal weights over time for regime-scheme 2. The optimal allocation changes as new returns data is incorporated and the regime-identifying variables are updated.

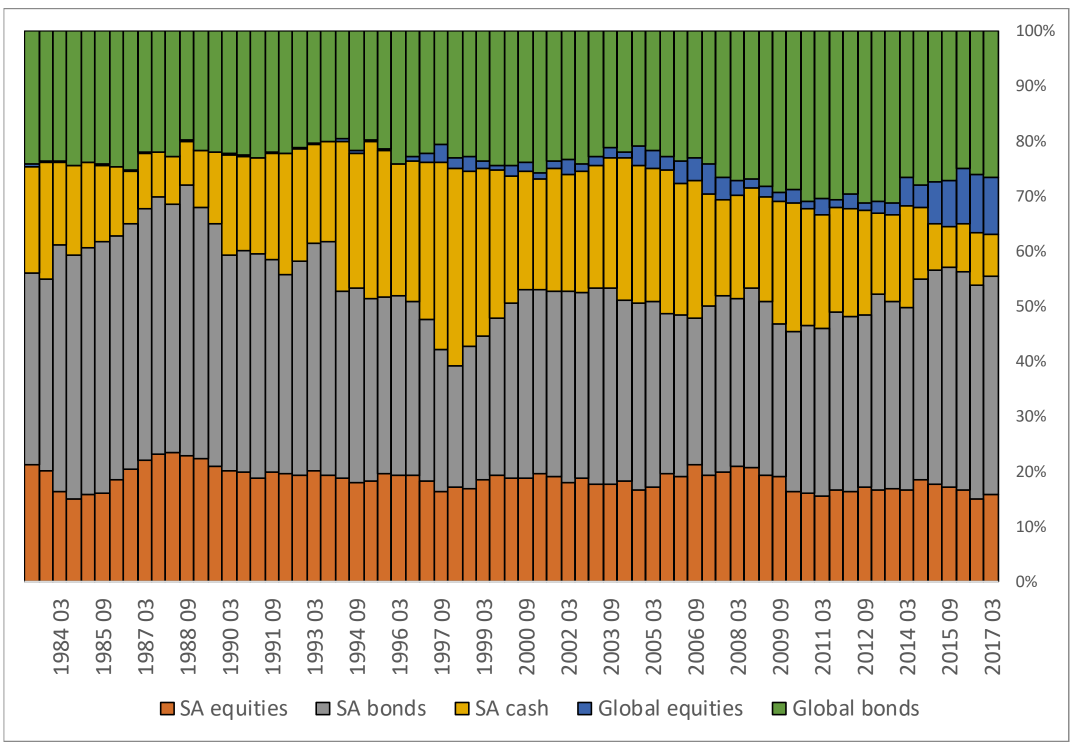

Figure 45: Optimal weights over time for regime-scheme 3. The optimal allocation changes as new returns data is incorporated and the regime-identifying variables are updated.

Table 24 below shows the average (over time) optimal weights of each regime-scheme. It reveals that, compared with the regime-ignorant optimal portfolios (MVO), the regime-cognisant optimal portfolios have all been consistently and materially underweight SA equities and overweight both SA bonds and SA cash from the start of the analysis.

The consistent underweight to SA equities in the regime-cognisant portfolios relative to MVO seems odd at first glance. However, recall that the optimisation finds the lowest standard deviation portfolio that meets the required real return – the reduction in equity thus reflects the fact that the monthly standard deviation can be reduced, while still (hopefully) meeting the required real return.

	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore
no regimes (MVO or SAA)	40%	9%	12%	4%	35%	43%	39%
regime-scheme 1	13%	29%	35%	6%	17%	19%	23%
regime-scheme 2	16%	33%	24%	4%	23%	20%	27%
regime-scheme 3	19%	35%	20%	2%	24%	21%	27%

Table 24: Average optimal allocation across all periods for MVO and the three regime-scheme strategies

5.2.2.1. Out-of-sample evaluation of the performance of optimal allocations

As we have assumed a 36-month investment horizon in the optimisation process, strictly speaking the most internally consistent evaluation of these optimal allocations would assume that the target asset allocation is left unchanged for a period of 36 months. Table 25 below shows the performance of each of the derived optimal portfolios for which 36-month returns had elapsed at the time of writing:

	36-month real returns				36-month annualised monthly standard deviation				36-month Sharpe ratios			
	regimes-ignorant (i.e. MVO)	regime-scheme 1	regime-scheme 2	regime-scheme 3	regimes-ignorant (i.e. MVO)	regime-scheme 1	regime-scheme 2	regime-scheme 3	regimes-ignorant (i.e. MVO)	regime-scheme 1	regime-scheme 2	regime-scheme 3
2014 Mar	1.9%	2.6%	2.6%	2.4%	5.6%	3.4%	4.1%	3.8%	0.17	0.49	0.42	0.39
2013 Sep	4.9%	3.7%	3.9%	4.0%	5.6%	3.4%	3.7%	3.8%	0.80	0.94	0.92	0.92
2013 Mar	7.9%	4.5%	5.2%	5.0%	5.9%	3.9%	4.0%	4.3%	1.31	1.11	1.24	1.12
2012 Sep	8.5%	4.8%	5.1%	5.4%	6.2%	3.7%	3.9%	4.1%	1.36	1.30	1.31	1.30
2012 Mar	10.9%	6.3%	7.2%	7.2%	6.3%	3.8%	3.8%	4.1%	1.72	1.64	1.86	1.77
2011 Sep	9.4%	5.2%	5.9%	6.0%	6.5%	3.7%	4.0%	4.0%	1.51	1.54	1.59	1.61
2011 Mar	9.6%	6.2%	6.9%	7.4%	6.3%	4.0%	4.2%	4.6%	1.62	1.71	1.81	1.77
2010 Sep	9.1%	5.4%	5.9%	6.0%	7.1%	4.2%	4.4%	4.8%	1.33	1.37	1.41	1.32
2010 Mar	7.6%	4.8%	5.6%	5.9%	5.6%	3.3%	3.7%	3.9%	1.31	1.38	1.44	1.45
2009 Sep	6.7%	4.9%	5.6%	5.7%	5.3%	3.4%	3.7%	4.0%	1.03	1.09	1.19	1.15
2009 Mar	5.6%	3.4%	4.1%	3.6%	6.4%	3.9%	4.2%	4.5%	0.68	0.53	0.66	0.51
2008 Sep	4.3%	4.3%	4.3%	4.7%	7.5%	4.9%	5.5%	5.7%	0.22	0.33	0.29	0.35
2008 Mar	-1.9%	0.2%	-0.1%	0.0%	8.8%	5.7%	6.2%	6.3%	-0.47	-0.35	-0.37	-0.36
2007 Sep	0.4%	0.8%	1.2%	1.3%	9.2%	6.0%	6.4%	6.6%	-0.19	-0.22	-0.15	-0.13
2007 Mar	-2.2%	-1.7%	-1.6%	-0.9%	9.3%	5.1%	5.8%	6.2%	-0.39	-0.61	-0.53	-0.38
2006 Sep	-1.2%	-0.9%	-0.2%	-0.1%	8.9%	5.3%	5.9%	6.1%	-0.32	-0.48	-0.30	-0.27
2006 Mar	1.4%	0.1%	1.0%	2.1%	8.7%	5.1%	5.6%	6.1%	0.01	-0.25	-0.07	0.11
2005 Sep	5.3%	2.2%	2.9%	3.4%	7.9%	4.1%	4.5%	4.9%	0.57	0.36	0.48	0.55
2005 Mar	14.5%	6.2%	7.9%	7.6%	7.6%	3.7%	4.1%	4.1%	1.69	1.24	1.51	1.45
2004 Sep	13.3%	7.5%	7.7%	7.7%	7.0%	3.2%	3.9%	3.9%	1.58	1.68	1.41	1.45
2004 Mar	16.5%	9.1%	9.8%	10.1%	7.5%	3.6%	4.2%	4.4%	1.78	1.67	1.61	1.59
2003 Sep	15.6%	8.0%	8.6%	9.4%	9.1%	4.3%	5.1%	5.0%	1.41	1.21	1.13	1.29
2003 Mar	14.1%	8.4%	9.2%	9.0%	11.0%	5.4%	6.3%	6.5%	0.89	0.75	0.77	0.72
2002 Sep	3.9%	4.1%	5.2%	4.8%	12.9%	6.5%	7.0%	7.6%	-0.06	-0.09	0.07	0.01
2002 Mar	-1.7%	1.7%	2.1%	2.2%	12.3%	6.5%	7.3%	7.3%	-0.50	-0.42	-0.32	-0.30
2001 Sep	1.7%	2.8%	2.7%	3.6%	14.7%	7.2%	8.2%	7.9%	-0.15	-0.16	-0.15	-0.04
2001 Mar	2.6%	3.0%	4.4%	3.9%	13.7%	7.9%	7.6%	8.6%	-0.09	-0.10	0.08	0.02
2000 Sep	2.5%	3.5%	5.2%	5.2%	13.0%	7.2%	7.3%	7.7%	-0.11	-0.05	0.18	0.17
2000 Mar	1.9%	3.3%	4.6%	4.5%	11.2%	5.8%	6.1%	6.5%	-0.07	0.12	0.32	0.29
1999 Sep	9.4%	9.0%	10.2%	9.9%	11.1%	5.9%	6.1%	6.3%	0.63	1.12	1.29	1.19
1999 Mar	15.6%	10.8%	11.5%	12.0%	10.3%	5.5%	5.9%	6.1%	1.09	1.15	1.21	1.23
1998 Sep	14.0%	12.9%	14.3%	14.6%	8.9%	4.2%	4.6%	5.1%	0.73	1.26	1.47	1.39
1998 Mar	8.5%	9.1%	9.3%	9.2%	11.0%	5.3%	6.0%	6.2%	0.12	0.36	0.35	0.33
1997 Sep	11.6%	10.7%	10.3%	10.4%	11.8%	4.6%	6.4%	6.3%	0.27	0.50	0.28	0.31
1997 Mar	10.9%	12.3%	11.0%	11.3%	10.5%	5.4%	6.3%	7.0%	0.10	0.45	0.17	0.20
1996 Sep	6.5%	10.0%	8.8%	8.6%	10.9%	6.0%	6.4%	7.8%	-0.34	-0.05	-0.23	-0.21
1996 Mar	6.7%	8.0%	7.9%	8.5%	10.4%	7.5%	7.3%	7.9%	-0.24	-0.16	-0.17	-0.08
1995 Sep	4.6%	6.8%	5.4%	5.3%	9.7%	5.9%	7.1%	7.6%	-0.28	-0.09	-0.27	-0.26
1995 Mar	10.3%	12.3%	11.5%	10.9%	5.8%	2.7%	3.1%	3.5%	0.30	1.34	0.93	0.66
1994 Sep	7.4%	10.0%	9.1%	9.4%	6.2%	2.9%	3.4%	3.7%	-0.01	0.86	0.51	0.54
1994 Mar	6.5%	5.8%	5.6%	5.7%	6.5%	3.6%	4.1%	4.1%	0.15	0.08	0.02	0.06
1993 Sep	11.1%	6.9%	7.2%	7.2%	7.7%	5.2%	5.9%	6.4%	0.79	0.35	0.37	0.33
1993 Mar	11.2%	7.7%	8.0%	8.5%	7.7%	5.7%	6.4%	6.8%	0.90	0.60	0.59	0.62
1992 Sep	10.5%	6.3%	6.8%	7.1%	7.8%	5.3%	5.9%	6.2%	0.86	0.47	0.51	0.52
1992 Mar	6.8%	5.1%	5.5%	5.7%	8.4%	4.9%	5.7%	5.8%	0.47	0.45	0.48	0.49
1991 Sep	7.8%	5.2%	5.7%	6.1%	8.0%	4.7%	5.4%	5.9%	0.67	0.58	0.60	0.61
1991 Mar	9.0%	6.5%	7.2%	8.1%	7.5%	3.8%	4.4%	4.9%	0.80	0.92	0.93	1.04
1990 Sep	5.8%	5.8%	6.9%	7.3%	7.3%	3.4%	4.0%	4.2%	0.36	0.75	0.94	0.97
1990 Mar	1.2%	3.5%	3.6%	3.4%	8.1%	3.6%	4.0%	4.3%	-0.36	-0.18	-0.12	-0.16
1989 Sep	-0.9%	3.5%	3.4%	3.2%	8.1%	3.4%	3.9%	4.5%	-0.62	-0.21	-0.19	-0.22
1989 Mar	1.8%	3.3%	2.9%	2.7%	8.0%	3.4%	4.1%	4.6%	-0.30	-0.28	-0.32	-0.32
1988 Sep	4.6%	3.2%	3.0%	3.1%	8.5%	3.7%	4.5%	5.1%	0.08	-0.23	-0.23	-0.17
1988 Mar	5.6%	4.3%	4.4%	4.8%	8.7%	3.9%	5.0%	5.7%	0.21	0.13	0.13	0.18
1987 Sep	-2.5%	-0.1%	-0.8%	-1.1%	11.7%	5.7%	6.8%	7.5%	-0.46	-0.53	-0.54	-0.53
1987 Mar	2.1%	1.0%	1.0%	1.3%	11.9%	5.4%	6.7%	7.3%	0.11	0.03	0.04	0.07
1986 Sep	1.3%	0.4%	0.5%	0.7%	11.4%	5.5%	6.1%	7.1%	0.20	0.24	0.23	0.23
1986 Mar	1.6%	0.5%	0.8%	1.6%	11.5%	5.5%	6.7%	7.7%	0.34	0.51	0.47	0.50
1985 Sep	-1.0%	-1.8%	-0.8%	-0.5%	11.0%	4.6%	6.2%	7.0%	0.21	0.34	0.41	0.41
1985 Mar	-2.3%	-1.9%	-2.3%	-1.7%	10.5%	5.1%	5.9%	6.7%	0.01	0.11	0.03	0.12
1984 Sep	4.8%	0.1%	0.2%	0.2%	8.3%	4.6%	6.0%	6.3%	0.74	0.31	0.25	0.25
1984 Mar	-0.3%	-1.5%	-2.7%	-2.1%	7.6%	5.7%	6.5%	7.8%	-0.07	-0.31	-0.46	-0.30
1983 Sep	2.1%	0.2%	-1.5%	-0.6%	10.3%	6.4%	7.5%	7.4%	-0.05	-0.37	-0.54	-0.44
1983 Mar	-0.1%	-1.5%	-4.2%	-3.7%	8.7%	5.7%	6.2%	6.7%	-0.46	-0.95	-1.33	-1.14
arithmetic average	5.8%	4.6%	4.8%	5.0%	8.9%	4.8%	5.4%	5.8%	0.41	0.43	0.44	0.45
% of instances regime beats no regime	n/a	37%	43%	46%	n/a	100%	100%	98%	n/a	52%	60%	62%

Table 25: 36-month performances of optimal strategies (63 rolling periods)

We can see in Table 25 that all four schemes underperformed the required real return, on average. The regime-based optimal allocations delivered on average lower 36-month returns than the regime-free scheme, but also experienced significantly lower standard deviations. Overall, the average Sharpe ratio increases with the incorporation of regimes.

We could evaluate the performance in a slightly more pragmatic but less theoretically consistent manner by assuming the point of view of an asset manager who re-evaluates his portfolio positioning continuously over time. He may still take a 36-month view in how he positions his tactical tilts, but nevertheless adjusts these tilts as new information becomes available.

From this point of view a single cumulative return and single standard deviation can be calculated for every regime setup for the full period (1982 to 2017). Table 26 depicts these returns and standard deviations, along with the resultant Sharpe ratio, as well as the tracking error of each regime-scheme to no regime (or MVO):

	no regime	regime 1	regime 2	regime 3
return	13.79%	12.98%	13.32%	13.42%
monthly standard deviation	8.91%	4.54%	5.18%	5.52%
real return	5.64%	4.83%	5.16%	5.26%
Sharpe ratio	0.32	0.46	0.47	0.45
tracking error with no regime	0.00%	5.30%	4.76%	4.54%

Table 26: Overall performance of optimal strategies with 6-monthly updating of target allocation. Here the target allocation changes every 6 months, in spite of the fact that they are optimised for 36-month investment horizons.

The results in Table 26 are broadly similar to those of Table 25. All four methodologies fell slightly short of the target real return of 6%. However, in risk terms, all three regime-schemes once again outperformed the regime-ignorant scheme materially, and as a result produced easily superior risk-adjusted returns.

In order to frame the improvement in the Sharpe ratio seen above purely in terms of returns (to obtain a sense of the enhancement from a pure return point of view), consider the resultant statistics, depicted in Table 27, when we combine each of the four schemes with a cash portfolio so that the resultant strategies all have the same overall monthly standard deviation:

	no regime	regime 1	regime 2	regime 3
weight to additional cash component	**50%**	**0%**	**13%**	**19%**
return	**12.45%**	**12.98%**	**13.02%**	**12.97%**
monthly standard deviation	**4.54%**	**4.54%**	**4.54%**	**4.54%**
real return	**4.30%**	**4.83%**	**4.86%**	**4.81%**
Sharpe ratio	**0.34**	**0.46**	**0.47**	**0.45**
tracking error with no regime	**0.00%**	**2.02%**	**1.88%**	**1.88%**
information ratio relative to no regime	**n/a**	**0.26**	**0.30**	**0.28**

Table 27: Overall performance of rolling optimal portfolios controlled for monthly standard deviation. An additional cash component is added in order to control for monthly standard deviation, enabling the quantification of the pure gains in return made by investing in regime-cognisant strategies. Regime-scheme 1 experienced the lowest standard deviation, hence additional cash is added to all the other strategies.

Regime-scheme 1 experienced the lowest standard deviation, hence additional cash is added (the top row of Table 27) to all the other strategies until their overall monthly standard deviation is in line with that strategy. The average asset allocation in these risk-controlled strategies (Table 28), is quite telling:

	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore
no regimes (MVO or SAA)	20%	5%	56%	2%	18%	22%	19%
regime-scheme 1	13%	29%	35%	6%	17%	19%	23%
regime-scheme 2	14%	29%	34%	3%	20%	17%	23%
regime-scheme 3	15%	29%	35%	2%	20%	17%	22%

Table 28: Average optimal allocation across all periods in the risk-controlled strategies with additional cash component

The striking similarity between the three regime-cognisant portfolios evident from Table 28 suggests that the main effect of the three regime-changing mechanisms relative to each other is an adjustment of tilt between SA cash and a portfolio of risky assets. This is further corroborated by the quite low tracking errors between these three strategies: the annual tracking errors between the three risk-adjusted regime-cognisant strategies are between 0.61% and 0.82% (though an increase in the tactical pressure would likely increase these tracking errors). Intuitively it makes sense to allocate more to cash if there is greater uncertainty about the regime (comparable to the hedging and myopic components of Merton, 1969) – the certainty about the prevailing regime is, in essence, the difference between these three mechanisms.

Another striking result in Table 28 is the relatively high allocation to SA bonds in the regime-cognisant portfolios (while having similar allocations in total to bonds and cash across all four strategies). The conclusion one can draw is that bonds become relatively more attractive than cash when regimes are incorporated. This is also broadly consistent with our discussions and the graphs in Section 4.1.4.2 comparing the long-term and short-term volatility of these two asset classes.

Consistent with low tracking errors between the three risk-controlled regime-cognisant portfolios, the cumulative returns for these risk-controlled strategies are very similar.

This is displayed in Figure 46, which depicts the growth of R1 invested in 1985 for each of these risk-controlled strategies:

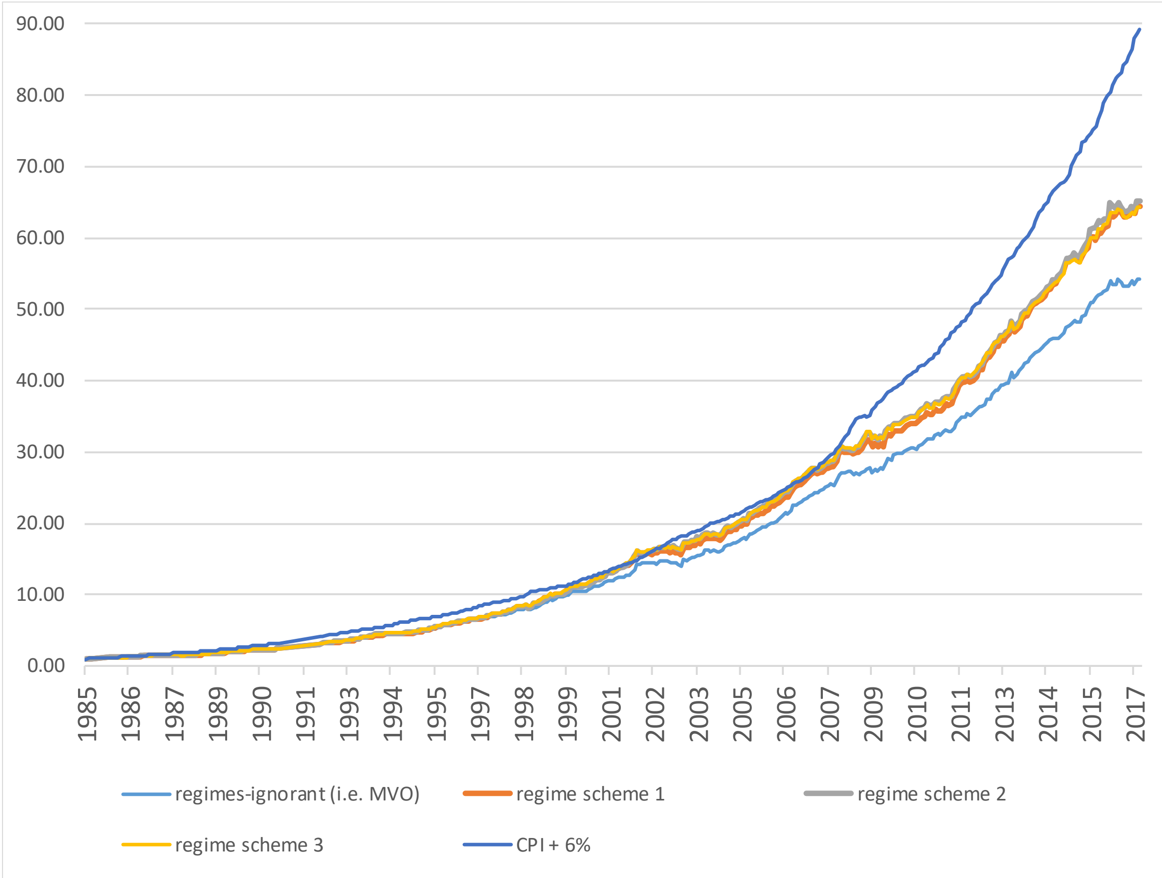

Figure 46: Growth of R1 invested in 1985 for the risk-controlled strategies with additional cash component. When risk in the form of monthly standard deviation is controlled, the regime-cognisant strategies exhibit higher returns than traditional MVO.

5.2.3. Regime-cognisant optimal portfolios of other risk criteria and block lengths (including out-of-sample backtests)

So far in Section 5.2, we have focussed on an arbitrary block length of 12 months, and only considered the monthly standard deviation criterion. However, as we have shown in Section 4.1.4.2, the standard deviation of asset returns depend on periodicity, and thus it can be argued that monthly standard deviation is not a strictly valid risk criterion for an investor with an investment horizon longer than one month.

In this section we will attempt to address these two shortcomings. We will test three block lengths (one, 12 and 24 months) that are broadly representative of the possible range of values the block length could take on, for three different risk criteria (monthly standard deviation, standard deviation over horizon and average shortfall). Once again we employ data from 1972 to 2017 to derive optimal portfolios at six month intervals from 1982 to 2017. This means that all optimal portfolios are based on at least 10 years of monthly returns and regime-identifying data.

The detailed results are shown in Appendix 7.4. For every criterion, there are 12 distinct optimal portfolios, representing every possible combination of the four possible regime-schemes (no regime, regime-scheme 1, regime-scheme 2, and regime-scheme 3) and the three different block lengths (denoted respectively as “bl1”, “bl12”, and “bl24”). If we rank the out-of-sample historical performance of each of these possible portfolios, for the case where we update the asset allocation every six months and the case where we do not update the asset allocation for the entire investment horizon, we arrive at Table 29 and Table 30:

	update every six months		
	shortfall	mth stdev (ranked by SR)	sd horizon (ranked by SR wrt sd horizon)
1	shortfall bl24 / regime 3	mth stdev bl12 / regime 2	sd horizon bl24 / no regime
2	shortfall bl24 / regime 2	mth stdev 24 / regime 1	sd horizon bl24 / regime 3
3	shortfall bl1 / regime 3	mth stdev bl12 / regime 1	sd horizon bl24 / regime 2
4	shortfall bl12 / regime 3	mth stdev bl12 / regime 3	sd horizon bl24 / regime 1
5	shortfall bl24 / regime 1	mth stdev 24 / regime 3	sd horizon bl12 / no regime
6	shortfall bl12 / regime 2	mth stdev 24 / regime 2	sd horizon bl12 / regime 3
7	shortfall bl12 / regime 1	mth stdev bl1 / regime 3	sd horizon bl12 / regime 2
8	shortfall bl12 / no regime	mth stdev bl1 / regime 1	sd horizon bl1 / no regime
9	shortfall bl24 / no regime	mth stdev bl1 / regime 2	sd horizon bl1 / regime 3
10	shortfall bl1 / no regime	mth stdev bl1 / no regime	sd horizon bl12 / regime 1
11	shortfall bl1 / regime 2	mth stdev 24 / no regime	sd horizon bl1 / regime 2
12	shortfall bl1 / regime 1	mth stdev bl12 / no regime	sd horizon bl1 / regime 1

Table 29: Ranked from best to worst for the case where the asset allocation is updated every six months (one calculation spanning the entire the period 1982 to 2017)

	don't update every six months		
	shortfall	mth stdev (ranked by SR)	sd horizon (ranked by SR wrt sd horizon)
1	shortfall bl24 / regime 3	mth stdev 24 / regime 2	sd horizon bl24 / no regime
2	shortfall bl24 / regime 2	mth stdev 24 / regime 3	sd horizon bl12 / no regime
3	shortfall bl24 / regime 1	mth stdev bl1 / regime 3	sd horizon bl1 / no regime
4	shortfall bl24 / no regime	mth stdev bl1 / regime 2	sd horizon bl24 / regime 3
5	shortfall bl12 / no regime	mth stdev bl1 / no regime	sd horizon bl24 / regime 2
6	shortfall bl12 / regime 3	mth stdev bl12 / regime 3	sd horizon bl24 / regime 1
7	shortfall bl1 / regime 3	mth stdev bl12 / regime 2	sd horizon bl12 / regime 3
8	shortfall bl1 / no regime	mth stdev bl12 / no regime	sd horizon bl12 / regime 2
9	shortfall bl12 / regime 2	mth stdev 24 / regime 1	sd horizon bl1 / regime 3
10	shortfall bl12 / regime 1	mth stdev bl12 / regime 1	sd horizon bl12 / regime 1
11	shortfall bl1 / regime 2	mth stdev 24 / no regime	sd horizon bl1 / regime 2
12	shortfall bl1 / regime 1	mth stdev bl1 / regime 1	sd horizon bl1 / regime 1

Table 30: Ranked from best to worst for the case where the asset allocation is updated only at the beginning of each 36-month investment horizon (the average of the result of 63 rolling 36-month periods between 1982 and 2017)

The rankings in Table 29 and Table 30 are from best (1) to worst (12). Note that in the case of monthly standard deviation and standard deviation over horizon, *we rank the* *results according to the Sharpe ratio,* employing the relevant measure of risk in each case in the denominator.

To further simplify things, Table 31 below shows the average ranking across all three criteria in the tables above, within various categories:

	average rank		
	update AA	don't update AA	both
no regime	8.2	5.2	6.7
regime 1	7.0	9.3	8.2
regime 2	6.2	6.4	6.3
regime 3	4.6	5.0	4.8
bl=1	9.2	8.1	8.6
bl=12	6.1	7.3	6.7
bl=24	4.3	4.1	4.2
bl=1 / no regime	9.3	5.3	7.3
bl=1 / regime 1	8.5	9.3	8.9
bl=1 / regime 2	10.3	8.7	9.5
bl=1 / regime 3	6.3	6.3	6.3
bl=12 / no regime	8.3	5.0	6.7
bl=12 / regime 1	6.7	10.0	8.3
bl=12 / regime 2	4.7	8.0	6.3
bl=12 / regime 3	4.7	6.3	5.5
bl=24 / no regime	7.0	5.3	6.2
bl=24 / regime 1	3.7	6.0	4.8
bl=24 / regime 2	3.7	2.7	3.2
bl=24 / regime 3	2.7	2.3	2.5

Table 31: Average rank across all three risk criteria of various subgroups

On the basis of insights gleaned by studying Table 29, Table 30 and Table 31, we are now in a position to make the following general statements:

- On average, the regime-scheme 3 (i.e. state changes occur as in chronological history), outperforms the other regime-schemes as well as no regime-scheme, followed by regime-scheme 2, and then no regime-scheme[21].
- Longer block lengths generally perform better than shorter block lengths.
- It is therefore no surprise that regime-scheme 3 in combination with a block length of 24 months performed the best overall, on average.
- The results are broadly consistent for both updating and not updating the asset allocation every six months

Though the possibility that the relatively small improvements in performance due to employing block lengths longer than one month and regime-schemes may be due to

[21]However, within the most successful block length, namely 24, regime-scheme 1 outperforms the no regime-scheme.

randomness cannot be ruled out, these results provide preliminary evidence in favour of these techniques.

Compared with classic MVO (i.e. a block length of one month and no regime-scheme), the approaches that employ block lengths greater than one month and optimise specifically to the risk criterion in question tend to yield superior performance (evaluated on the same basis as above). Table 32 and Table 33 below indicate whether the optimal portfolio in each case outperformed classic MVO, for a block length of 12 months and 24 months, respectively:

	update AA every six months			don't update AA every six months			
	shortfall	mth stdev (SR)	sd horizon (SR)	shortfall	mth stdev (SR)	sd horizon (SR)	
	4 out of 4	3 out of 4	3 out of 4	2 out of 4	3 out of 4	1 out of 4	16 out of 24
no regime	1	0	1	1	0	1	4 out of 6
regime-scheme 1	1	1	0	0	1	0	3 out of 6
regime-scheme 2	1	1	1	0	1	0	4 out of 6
regime-scheme 3	1	1	1	1	1	0	5 out of 6

*Table 32: Comparing the risk-adjusted performance of the optimal portfolio with explicit reference to a specific risk criterion (**12 month block length**) to that of the classic MVO portfolio (1 indicates outperformance, 0 indicates underperformance)*

	update AA every six months			don't update AA every six months			
	shortfall	mth stdev (SR)	sd horizon (SR)	shortfall	mth stdev (SR)	sd horizon (SR)	
	4 out of 4	3 out of 4	4 out of 4	4 out of 4	2 out of 4	1 out of 4	18 out of 24
no regime	1	0	1	1	0	1	4 out of 6
regime-scheme 1	1	1	1	1	0	0	4 out of 6
regime-scheme 2	1	1	1	1	1	0	5 out of 6
regime-scheme 3	1	1	1	1	1	0	5 out of 6

*Table 33: Comparing the risk-adjusted performance of the optimal portfolio with explicit reference to a specific risk criterion (**24 month block length**) to that of the classic MVO portfolio (1 indicates outperformance, 0 indicates underperformance)*

It should be noted, however, that the improvements are relatively marginal and do not by themselves conclusively argue for the techniques employed. Table 34 and Table 35 below show how respectively the average shortfall portfolio and standard deviation over horizon portfolios (with a block length of 24 months) compare with the MVO portfolios (with block lengths of one month and 24 months):

		avg shortfall			
		no regime	regime 1	regime 2	regime 3
Update AA every 6 months	MVO portfolio (bl=1)	2.0%	2.2%	2.1%	1.6%
	MVO portfolio (bl=24)	1.9%	1.6%	1.4%	1.4%
	avg shortfall portfolio (bl=24)	1.8%	1.5%	1.4%	1.4%
Don't update AA every 6 months	MVO portfolio (bl=1)	2.1%	2.9%	2.4%	2.2%
	MVO portfolio (bl=24)	2.3%	2.3%	2.1%	2.0%
	avg shortfall portfolio (bl=24)	2.1%	2.0%	1.9%	1.9%

Table 34: The out-of-sample average shortfall of the optimal average shortfall portfolio (block length of 24) versus that of the MVO portfolios (with respective block lengths of 1 month and 24 months)

		Sharpe ratio with sd horizon			
		no regime	regime 1	regime 2	regime 3
Update AA every 6 months	MVO portfolio (bl=1)	61.7%	36.2%	37.9%	61.5%
	MVO portfolio (bl=24)	60.3%	65.0%	74.5%	69.8%
	sd horizon portfolio (bl=24)	73.4%	65.9%	70.2%	72.0%
Don't update AA every 6 months	MVO portfolio (bl=1)	51.3%	14.2%	30.4%	42.4%
	MVO portfolio (bl=24)	48.2%	37.8%	49.8%	48.8%
	sd horizon portfolio (bl=24)	58.1%	47.1%	48.5%	49.7%

Table 35: The out-of-sample Sharpe ratio (with sd horizon on the denominator) of the optimal sd horizon portfolio (block length of 24) versus that of the MVO portfolios (with respective block lengths of 1 month and 24 months)

It should be borne in mind that the block length could be further fine-tuned, which may yield superior improvements to those seen in Table 34 and Table 35. Similarly, our choice of the value of 10 for the tactical pressure parameter was arbitrary and static. A further possible avenue for improving the out-of-sample performance is a more sophisticated analytical approach to choosing the tactical pressure parameter at every point in time. Additionally, it is plausible that the benefits of our methodology are increasing in the length of the investment horizon. Finally, as we only backtested a single required real return (6%), further research is required to ascertain whether the gains may be more (or less) pronounced for other required returns.

6. CONCLUSIONS AND FURTHER RESEARCH

The higher level aim of this work was to demonstrate how the optimal allocation with respect to almost any risk measure can be derived under very general conditions with very few assumptions. Against this backdrop we set out to find the optimal strategic (regime-ignorant) and tactical (regime-cognisant) asset allocations for South African investors targeting a real return over a specific investment horizon by employing data-intensive, nonparametric methods. Various measures of risk were considered.

In order to account for possible non-normality of the joint distribution (regardless of whether conditional or unconditional on regime), we modelled the returns of each portfolio of assets by its own unique empirical distribution, employing returns from 1972 to 2017. We designed a bespoke genetic algorithm able to find the optimal portfolios under these very general conditions. In some cases (namely monthly standard deviation and standard deviation over investment horizon), our optimal portfolios could have been found via MVO methods, but in others (downside deviation, expected shortfall, maximum drawdowns) we could not have derived optimal portfolios by classical means (at least not without resorting to the assumption that returns are normally distributed), showcasing the potential usefulness of genetic algorithms in the context of portfolio optimisation.

The investment horizon is incorporated in two ways: firstly by employing risk criteria that directly make reference to investment horizon in their definitions (e.g. the average shortfall over the investment horizon and the probability of achieving the required return over the investment horizon), and secondly by capturing the time-dependencies by increasing the periodicity of asset returns.

The time-dependencies of asset class returns – chiefly short-term momentum and long-term mean reversion of SA equities returns – and the resultant importance of the choice of periodicity for the empirical distribution (i.e. the “block length”) were discussed in some detail. While a block length of one month minimises the standard errors of estimating the variance of SA equities (the most important asset class for a South African investor), a block length equal to the length of the investment horizon yields unbiased estimates. Although we employ a random block length to reduce the effect of choice of block length, the choice of expected block length remains the most contentious and important parameter of this analysis. As a result we resolved to examine the effect of this parameter on the outcome of optimal portfolio weights.

Both a regime-ignorant joint distribution (relevant to the choice of a long-term or strategic asset allocation) and a regime-cognisant distribution (relevant to the choice of a short-term or tactical asset allocation) were considered. For each case we found the optimal portfolios for not only the typical risk measure of standard deviation of monthly returns, but also the various investment horizon-dependent risk criteria.

The regime-ignorant optimal portfolios represent the portfolios that, in essence, performed the best over the period 1972 to 2017 in each of the respective problem settings. We believe the optimal portfolios tables for the intermediate block length, the one-month block length, the empirical block lengths and the insights gained by various carefully chosen comparisons are of significant practical value for investors, asset managers, and financial planning professionals seeking strategic asset allocation insights incorporating the oft-ignored time-dependencies of asset returns. Some highlights of the insights include estimates of the bias in the optimal weights produced by traditional MVO methods, estimates of the probabilities of achieving real return targets, and the relationship between investment horizon on the one hand and optimal weights, real return, and monthly standard deviation on the other.

A significant stumbling block of our approach is the derivation of optimal portfolios in this very general setting. Since the usual analytical methods employed in mean-variance optimisation to find the optimal portfolios cannot be used in this setting, we devised a bespoke genetic algorithm – a sophisticated metaheuristic approach – to help find the optimal portfolios for each problem setting.

The minimum monthly standard deviation (i.e. the classic MVO portfolio) performed surprisingly well in terms of limiting the maximum drawdowns. The optimal maximum drawdown portfolios found by the genetic algorithm performed only marginally better than the MVO portfolio, suggesting that the MVO portfolio is a good proxy for minimising drawdowns. Nevertheless, the optimal maximum drawdown portfolios were noted to have consistently higher exposure to local and global bonds, less exposure to SA cash (surprisingly), and also less exposure to global equities (see Table 20 on page 101).

The portfolio that maximises the probability of achieving the real return target (arguably the portfolio most relevant to the survival of asset managers) was, not surprisingly, found to be the most aggressive and risky portfolio in terms of equity allocation and monthly standard deviation. However, what was somewhat surprising was that these portfolios were not found to be uniformly increasingly aggressive for longer time horizons. However, the probability of achieving the target was clearly increasing in the investment horizon.

The average shortfall optimal portfolio is arguably (along with downside deviation over horizon) the most relevant risk measure to an investor with an explicit and analytically derived required real return that matches future assets and liabilities. As one would expect given its relatively lenient penalty on downside and no penalty on upside variation, the average shortfall over the horizon portfolio is, after the optimal portfolio for probability of achieving the target, the most aggressive and volatile of the optimal portfolios belonging to the various criteria. Also, the longer the investment horizon, the more aggressive and volatile the optimal portfolio. The optimal total equity weight for an investment horizon between 60 months and 120 months is between 26% and 45% higher than the typical MVO portfolio that targets the required real return (see Table 18 on page 98). This suggests that typical approaches may leave a long-term investor severely underexposed to equities. The downside deviation portfolio exhibits the same phenomena (14% to 37% higher optimal equity weight than MVO for investment horizons between 60 and 120 months), but slightly toned down due to the harsher penalty on the downside (see Table 18 on page 98).

The minimum standard deviation over the investment horizon portfolio is distinct from the MVO portfolio based on monthly data, due to the time-dependencies of returns. As a result, despite the fact that it *does* penalise downside variation, it still generally results in more aggressive and volatile portfolios as the investment horizon is increased (due mostly to long-run mean reversion in SA equities), albeit less so than the optimal average shortfall and downside deviation portfolios. The optimal equity allocation is between 4% and 8% higher than classic MVO for investment horizons between 60 months and 240 months (see Table 19 on page 99).

By comparing the equity allocation of the minimum standard deviation over horizon portfolio, the MVO portfolio (based on monthly returns), and the average shortfall portfolio, we could deduce that that the majority of the increase in the equity exposure of the latter is due to the lack of penalty on upside variation, with the remainder due to the time-dependencies of returns (see Figure 35 on page 108).

We compare each of our optimal portfolios with (generally) two distinct MVO portfolios: the MVO portfolio that satisfies the required return, and the MVO portfolio with the same level of return as the bespoke optimal portfolio (see Table 20, Table 21 and Table 22 from page 101). Our optimal portfolios generally improve upon both of these MVO portfolios with regards to the criterion being minimised, though much more significantly in the former case. All considered, results suggest that the common financial planning approach of calculating a "required real return" by simply solving the internal rate of return that balances an investor's assets and liabilities, and then recommending a fund that targets this real return, is materially suboptimal.

The fact that our bespoke optimal portfolios are materially different in terms of their asset weights from the MVO portfolio of the same return, proves that they are not merely points higher on the efficient frontier, but are in fact not on the traditional efficient frontier at all. The most notable and consistent pattern is that local equities and global bonds receive higher allocations (at the expense of global equities) and that the total offshore allocation reduces overall. Another interesting result is that, for the standard deviation over horizon portfolio, SA bonds receive significantly higher allocation than SA cash, likely due to the fact that the variability over longer periodicities of these two asset classes are very similar, and thus SA bonds, with the higher overall return, tends to dominate SA cash. Importantly, this calls into question whether traditional MVO portfolios over-allocates to cash and whether there truly is a place for SA cash in a strategic asset allocation or for an investor with a long investment horizon.

We also derived and discussed the empirically optimal portfolios (see Appendix 7.2.2 on page 143). Here the block length is fixed to be equal to the investment horizon. In other words, these optimal portfolios represent the portfolios that *actually* performed the best over the given criterion and the given horizon length, in history. As estimates of the optimal portfolios going forward, the empirically optimal portfolios effectively employ an unbiased estimate of the underlying distribution, but at the cost of higher standard errors, generally.

We then compared the optimal weights for the case of a block length of one month against the optimal weights for the intermediate block length (see Table 23 on page

110). This comparison is another estimate of the effect of the time-dependencies of asset returns on optimal allocations, which is seen to be material.

In the eight asset case (in Appendix 7.3 on page 147) we included local and offshore property and offshore cash in our optimisation. Global cash received practically no weight at all, suggesting that this asset class has no long-term, strategic role to play in an asset allocation. While local and global property received significant allocations, their optimal weights were decreasing in the investment horizon. We speculated that this may be due to the fact that they are increasingly dominated by their equity counterparts, which have very similar risk and correlation characteristics but yet higher overall returns. This finding suggests that the optimal allocation to property may be smaller than those found by studies that assume monthly returns are i.i.d.

Lastly, we incorporated regimes into our modelling to arrive at regime-cognisant optimal allocations. We did not classify each point in time as belonging to one of a finite set of regimes, but rather calculated the multivariate distance (over a collection of both economic and market variables) between each and every point in time.

Our regime classification greatly reduces the risk of misclassification, but this comes with a loss of ease of interpretability. We considered a static state as well as changing state regime mechanisms, and compared the results to standard MVO. Backtesting our three regime-cognisant optimal portfolios derived with data from 1972 to 2017 yielded results that exhibited quite consistent monthly standard deviation reduction, but with a sacrifice in overall returns. However, Sharpe ratios showed a clear improvement, and when controlled for risk, the return-enhancement was in the region of 0.5% to 0.6% per annum (see Table 27 on page 121). The three regime-changing mechanisms were noted to mostly express themselves through different exposures to SA cash. This is consistent with the fact that, in essence, each of these mechanisms represents a different level of certainty about the prevailing regime across the investment horizon.

We then widened our focus for the regime-dependent results to include two additional block lengths (one month and 24 months) as well as two additional risk criteria (average shortfall over horizon and standard deviation over horizon). We find that on average among these risk criteria, longer block lengths and the incorporation of regimes tended to outperform (on a risk-adjusted basis) optimal portfolios derived with a block length of one month and no regime-scheme (see Table 29 to Table 35 from page 123). A block length of 24 months and regime-scheme 3 – which, after an initial jump, changes states in accordance with chronological history – performed the best out-of-sample on average across these three risk criteria. Further, employing block lengths greater than one month tended to outperform the optimal portfolios derived with classic MVO (i.e. a block length of one month and ignoring regimes), though the effect sizes were generally small.

Further research

On a higher level this research shows how optimal portfolios can be found in a non-normal setting where each and every portfolio has a unique distribution based on its own historical returns. Further research could improve on the distributional assumptions by employing more sophisticated kernels than that of the empirical distribution. Such kernels could smooth out the distributions to remove some of the variance in their estimates.

Our choice of the “tactical pressure”, which determines the relative weight given to the returns belonging to any given point in time given the prevailing state, was arbitrary. One method that could potentially be used to decide on the value of this variable is cross-validation against some measure of returns or risk-adjusted returns.

The “optimal” block length is a nuisance parameter whose selection requires further research. Although a block length of one month appears to minimise the standard errors in the case of SA equities, it introduces sizeable bias. Out of the three block lengths tested (in the context of a 36-month investment horizon) out-of-sample, namely one month, 12 months and 24 months, the latter performed the best, on average. Although we have found some evidence that the optimal block length may be in the region of 24 months in the case of 36-month investment horizon, a more precise optimal block length and its relationship to the investment horizon remains an open question.

The magnitude of improvements in out-of-sample tests were marginal, and further experimentation with block lengths and investment horizons would be needed to confirm whether the use of block bootstrap in combination with our regime-identification methods represent a material improvement in the modelling of returns distributions that account for investment horizons.

In both the regime-ignorant and regime-cognisant scenarios, we assumed monthly rebalancing to a target allocation. We discussed the limitations of this assumption, and the requirements of the setting to render our analysis relevant and valid. In the case of regime-cognisant portfolios, a more realistic model would allow a more dynamic evolution of the asset allocation (for example, via a set of rebalancing rules), and the incorporation of transaction fees would be an important aspect of this model. A more sophisticated genetic algorithm could again be used to solve this dynamic programming problem.

7. APPENDICES

7.1. Solving for the optimal portfolios: The genetic algorithm

This section introduces the reader to genetic algorithms and discusses the bespoke genetic algorithm devised and employed to find the optimal portfolio among all the possible portfolios.

7.1.1. The need for metaheuristics in our problem setting

Mean-variance optimisation is typically solved via quadratic optimisation. While this procedure naturally caters for linear constraints on portfolio weights, it cannot be used to optimise other risk measures. In the present analysis we are not limited to normal distributions and only variance as a measure of risk; instead, as already discussed, returns are modelled with a generalisation of the empirical distribution. In this general setting a far more flexible optimisation routine is required.

Even for relatively large increments in portfolios weights, the number of portfolios quickly becomes prohibitively large from a computational point of view. For example, for the eight asset class case, the total number of possible portfolios that can be defined with increments ranging from 2.5% to 100% are listed by Table 36:

increment	number of possible portfolios
100%	8
50%	36
25%	330
10%	19,448
5%	888,030
2.5%	62,891,499

Table 36: Number of possible portfolios for various increments in the eight asset class case

To make the meaning of an increment clearer, for the case of increments of 5%, the 888 030 possible portfolios can be retrieved via the following algorithmic progression, as shown in Table 37:

	SA equities	SA bonds	SA cash	SA property	Global equities	Global bonds	Global cash	Global property
1	100%	0%	0%	0%	0%	0%	0%	0%
2	95%	5%	0%	0%	0%	0%	0%	0%
3	95%	0%	5%	0%	0%	0%	0%	0%
.	.	.	.	.	.	.	.	.
.	95%	0%	0%	0%	0%	0%	0%	5%
.	90%	10%	0%	0%	0%	0%	0%	0%
.	90%	5%	5%	0%	0%	0%	0%	0%
.	.	.	.	.	.	.	.	.
888028	0%	0%	0%	0%	0%	5%	0%	95%
888029	0%	0%	0%	0%	0%	0%	5%	95%
888030	0%	0%	0%	0%	0%	0%	0%	100%

Table 37: All possible portfolios for 5% increment in the eight asset class case

A granularity of 2.5% or finer would arguably be the minimum requirement in the setting of optimal portfolios. Thus, at a bare minimum, assessing each possible portfolio individually in the eight asset class case would involve performing a bootstrap simulation with at least a 1 000 iterations on approximately 63 million portfolios. If we assume 1 000 iterations required one second of computing time, it would require in the region of two years to investigate all of the possible portfolios.

We need a method that performs a similar function in our empirical setting as quadratic optimisation performs in the MVO setting. In other words, we need a method to search the arbitrarily large solution space (all possible long-only portfolios) that are sufficiently close to globally optimal from a practical standpoint. This brings us to the field of metaheuristics and genetic algorithms.

7.1.2. Introduction to genetic algorithms

Genetic algorithms (GA) are a class of metaheuristic algorithms that mimic the process of natural evolution by continually and probabilistically improving successive generations of a population, where, as the name suggests, each generation is derived from the previous. The field of metaheuristic deals with problems where an exact solution is not possible due to the large number of possible combinations in the solution space, and the limited availability of computing power. Typically there is a trade-off between calculation time and the quality of the solution, and a sufficiently good solution (or the best solution within the given time limits) is sought.

It is difficult to pinpoint the first implementation, but various versions of an algorithm that mimics evolution are known to have existed from at least the 1960s. As early as 1950, Alan Turing proposed "learning machines" with similar mechanisms to those exhibited by natural evolution (Turing, 1950).

Genetic algorithms are often employed in portfolio optimisation when real world constraints or practicalities define a complicated objective function that cannot be solved by existing methods or analytically (Chen, 2002). For example, standard mean-variance optimisation, and the assumptions of normality and identically and independently distributed returns, reduce the problem to a quadratic optimisation problem and is easily solved, even under constraints of linear combinations of assets. However, if the constraints on normality are relaxed, returns have intertemporal dependencies, or risk is not simply defined as variance, the problem must be solved by other means.

We will not attempt to give a detailed literature review of this area, as the field is too large and varied and it is not deemed within the scope of the present research. We do not intend to showcase or discuss the state of the art of methods in this area. Instead we want to give a broad and light introduction to the method and implement a version that works sufficiently well for the problem at hand.

We briefly list just a few sources of information for the interested reader.

Leinweber & Arnott (1995) published one of the first journal articles employing genetic algorithms in the context of tactical asset allocation.

Cheung, Kong, Tang & De Montreal (1996) solve a multi-period stochastic optimisation problem in the context of changing financial markets by maximising a utility function with a genetic algorithm.

Mulvey, Rosenbaum & Shetty (1997) employ metaheuristic methods to address the long-term financial planning context when posed as a dynamic programming problem with decision rules. The metaheuristic searches for the optimal parameters of the decision rules.

Chang, Yang & Chang (2009) is an example of portfolio optimisation with genetic algorithms when various measures of risk are assumed: semi-variance, mean absolute deviation and variance with skewness.

7.1.3. The components of genetic algorithms

Genetic algorithms is a term used to describe a very broad and varied class of algorithms. However, most genetic algorithms will employ most if not all of the following components or operators:

- Fitness function: The objective function being optimised;
- Population of individuals: A number of candidate solutions (in our setting, portfolios or asset allocations);
- Chromosomes: Constituent components of the individuals, in this case individual asset weights;
- Selection (alluding to natural selection) of one or more of the individuals (portfolios);
- Crossover operation (or mating procedure): a procedure that results in a new individual (portfolios) with characteristics inherited from two or more existing individuals;

- Mutation operation, a small random alteration to one or more individuals (portfolios); and
- Gradual improvement of a population (an evolution across generations).

The next section discusses some of these terms in more detail.

7.1.3.1. The fitness function

The objective function in our case depends on the risk criterion being optimised and the required real return. We maximise

$$\Omega(\boldsymbol{w}) = -criterion(\boldsymbol{w}) - C * \max(0, req\ real\ return - portfolio\ real\ return(\boldsymbol{w})) \tag{64}$$

where $\boldsymbol{w}$ is the portfolio weights and C must be large enough to essentially render almost any portfolio with a return below the required threshold less fit than any portfolio above the required real return.

In other words, we minimise the sum of (i) the criterion and (ii) a severe penalty for not adhering to the required return. For portfolios with a return above the threshold the second term will be zero and only the first term, the risk criterion, is relevant. For portfolios with a return below the threshold, the second term will almost always carry the most weight. Thus, among these latter portfolios, the ones with the higher return will usually have the better objective function.

One option would have been to include a penalty for not adhering to the linear constraints on weights (e.g. no short selling) in the objective function, but, as noted below, our genetic algorithm instead filters out non-viable portfolios at the population generation step, the crossover step and the mutation step. In retrospect, the former solution would be considerably more elegant, but it is not clear whether it would outperform our scheme.

7.1.3.2. Population

The population size used in this research varied from 50 to 100. For each individual, uniform random variables are generated, one for each asset, and then divided by their sum to force summation to one. The individuals are filtered for real return and linear constraints on the asset weights, and regenerated if necessary until the required population is reached. The return filter prompts a bootstrap simulation if a regime-scheme is employed, since such a simulation is required in that case to assess the expected return.

To save time, in the event that a large number of population generation iterations are required to fill the population (due to the return and allocation constraints being restrictive or the particular regime results in a very low expected real return, making it very unlikely for a random portfolio to be within constraints), the algorithm will start to mutate already generated individuals that adhered to constraints and were already added to the population, as a shortcut to finding viable individuals.

7.1.3.3. Selection

Selection refers to the process of obtaining individuals that will procreate to produce offspring that will form the new population. The parent generation dies off as soon as a new generation is produced. As is the case in natural selection seen in nature, the more "fit" individuals (i.e. individuals with better objective functions) have a higher probability of being selected for mating.

Selecting parents for the next generation is a delicate balance between selecting sufficiently "good" individuals to ensure the population improves, while also allowing randomness to ensure that some diversity is retained in every successive generation. If we only selected the very best individuals, we would very quickly converge to a local optimum, and conversely, if the selection were completely random, successive populations would show very little improvement. The other requirement of selection is that it involves as few calculations as possible, to ensure speed.

There are several ways to perform selection.

A popular method is roulette wheel selection in which the probability of selecting an individual is directly proportional to its objective function value. It is an efficient method to select individuals when the objective function is positive and maximisation is sought. If the objective function is minimised or negative, it is less efficient as the objective function of each individual must first be transformed to a function that is positive and increasing in the fitness of the individual. As we are presently attempting to minimise various definitions of risk (of which the objective function is a decreasing function), we opted instead for "tournament selection".

Tournament selection involves repeatedly and randomly (with replacement) selecting two individuals, comparing their fitness, and then choosing the individual with the better objective function. This is repeated until we have selected as many individuals as is in the population (P), since, as will be explained below, one pair of individuals will each produce two offspring. Clearly, as there is selection with replacement, each individual may be selected between 0 and P times. Only pairwise comparisons are made, so relatively little calculation is required. Better solutions will have a higher probability of being selected, but it is also the case that relatively weak solutions can also be retained – any individual must only be more optimal than the one other individual it happens to be in competition with. This ensures that some diversity is maintained across generations.

7.1.3.4. The crossover operation

The crossover operation mimics the production of offspring between two parent organisms in a population. As in the natural world, the offspring will retain some of the chromosomes of both of the parents. The most appropriate crossover operation is highly dependent on the specific application. In our current setting, an obvious and convenient choice is to simply average the two parent portfolios to arrive at the offspring – each chromosome (or asset weight) will have a value that is the average of the two parents. A slightly more sophisticated approach would be to consider a weighting scheme, i.e. a weighted average of two parents. This is the approach followed in this genetic algorithm. Two variations are employed: (1) arithmetic crossover; and (2) heuristic crossover.

Arithmetic cross-over

For each pair of parent individuals a random number, say w, between 0 and 1 is generated. Two offspring are then produced: (1) by assigning weight of w and 1-w to each parent; and then (2) reversing weights to produce another. This weighting scheme is very simple and fast, but it exhibits no intelligence in the weighting mechanism. Apart from being simple, intuitive, and fast, another good characteristic of this operator in our current setting is that it if the two parents (portfolios) are "valid" (i.e. within the linear allocation constraints and with a return that meets the target return), the offspring will also always be valid.

Heuristic crossover

For each pair of individuals (portfolios), we again randomly generate a weight w between 0 and 1. However, with this method, which was taken from the work of Ackora-Prah, Gymaerah & Andam (2014), we first discriminate between the fitter and less fit parents (as measured by the objective function), and then produce two offspring: (1) an offspring that starts with the fitter parent and adjusts it away from the weaker parent in the direction of the fitter parent in the following manner:

$$offspring\ 1 = Best\ parent + w * (best\ parent - worst\ parent) \quad (65)$$

$$offspring\ 2 = Best\ parent \quad (66)$$

While this method uses some intelligence in the adjustment made in offspring 1, it also requires extra calculation in that the better parent in each pair must be established. There is also no guarantee that offspring 1 will be valid with regards to the linear allocation constraints on individual assets. As a result, since we did not include a penalty term in the objective function for not adhering to allocation constraints, we need to remove invalid portfolios and then repeat the heuristic crossover until we have the required number of offspring.

In the genetic algorithm half of offspring will be produced by arithmetic crossover and the other half by heuristic crossover.

7.1.3.5. Mutation operation

The mutation operation typically makes a random change to one or more chromosomes in typically a relatively small proportion of individuals, with two broad aims: (i) to maintain some diversity in the population; and (ii) to create the potential to arrive, from time to time, at fitter individuals with chromosomes potentially not already present in the population.

Mutation guarantees that there will always be at least some diversity, even if the population being mutated consists of only a proliferation of a single individual, thereby sustaining the potential to find fitter individuals through all generations. Put another way, it ensures there is always a possibility of escaping from a local optima, rendering all GA iterations potentially useful.

In our genetic algorithm, one third of the individuals selected for mutation will receive exactly one of the three possible mutation procedures:

(1) Randomly generate a normally distributed random variable with a mean of 0 and a standard deviation of 0.2 and add this number to randomly selected asset weight. Normalise the weights by dividing the entire portfolio by the sum of its weights to ensure the portfolio again adds up to one. The net effect is to maintain the proportions between all asset classes except the mutated asset class in the original individual, while changing the proportion between the selected asset class and the rest of the asset classes.
(2) Randomly generate a normally distributed number with a mean of 0 and a standard deviation of 0.01, and add and subtract this number to or from two randomly selected asset weights.
(3) Randomly generate two normally distributed random variable with a mean of 0 and a standard deviation 0.2, and add these two numbers to two randomly selected asset weights.

The third mutation procedure is the most disruptive, followed by the first procedure. The second procedure is by far the least disruptive, as the size of the change is small and only affects two asset weights, while maintaining the proportion of those two weights versus the other remainder of the asset weights.

All of these mutation methods suffer from the same drawback we saw in heuristic crossover: the resulting individuals will not necessarily fall within any potential linear constraints on asset weights, regardless of whether the original individual being mutated had. If the produced individual is not viable, it is discarded and the whole process is repeated until a viable individual is found. However, since we only mutate a relatively small proportion of the population, the computational cost is arguably acceptable.

7.1.3.6. Our genetic algorithm

Now that we have introduced the basic building blocks of genetic algorithms, we can present the full genetic algorithm employed in this research.

- *Randomly generate the population*
- ***WHILE** termination conditions are not met DO*
    - ***IF** this is not a special 'while' iteration DO*
        - *Perform bootstrap with standard number of bootstrap iterations*
    - ***ELSE** it is a special 'while' iteration*
        - *Perform bootstrap with higher number of bootstrap iterations*
        - *Record the best individual (i.e. or solution or asset weight) in the population and add to a collection of "best" individuals*
        - *Note whether termination conditions have been met*
        - *Write population and best allocation to csv*
    - ***END IF** (special/not special)*
    - ***IF** termination conditions are not met*
        - *Perform tournament selection*
        - *Perform crossover operation*
        - *Perform mutation operation*
    - ***END IF** (termination conditions)*
- ***END WHILE** (termination conditions)*
- *Perform one final population bootstrap with very large number of bootstrap iterations on final population*
- *Record the best individual (i.e. or solution or asset weight) in the population and add to the collection of best portfolios*
- *Perform bootstrap on the collection of best portfolios with very large number of bootstrap iterations*

Algorithm 4: The genetic algorithm

Key features of the algorithm

1. The termination condition will be true if any one of the following conditions are met:

    i. Certain number of generations completed;
    ii. Certain amount of time elapsed (the specific time would be determined by the time constraints of the user); or
    iii. A certain number of 'special while iterations' (see 2 below) are completed with no improvement in the average quality of the population (i.e. the average of all the objective function values across the population). This condition could, loosely speaking, be considered a convergence parameter.

2. The iterations of the while loop are divided into two types:

    - a non-special while iteration; and
    - a special while iteration

    The special iteration only occurs with a pre-specified frequency (typically set to between 15 and 50 in this research).

    In contrast to the non-special iteration, during a special iteration we check whether termination conditions have been met and record the population and the best solution found. By distinguishing between these two types of iterations we save material calculation time during the non-special iterations, which are the large majority of iterations.

    During the special iteration, we perform a significantly larger number of bootstrap iterations for two main reasons:

    i. To ensure that our convergence Boolean variable is not triggered due to randomness rather than true convergence; and
    ii. To ensure that the "best" portfolio found is actually the best with a reasonably high probability, so that we have a reasonably accurate record of the best portfolios found over time. This ensures that we do no lose good solutions over time due to randomness in the bootstrap resampling.

    In this research we typically set the non-special bootstrap iterations to 1 000, and special iterations to 5 000.

3. The allocation filter is applied at every point where new individuals are generated (be it random, mutated or via crossover). While time-consuming at first, this hugely narrows the search space, ensuring that time is not wasted on inadmissible portfolios, and possibly ultimately saving calculation time.

4. After convergence, one last bootstrap is performed with a very large number of iterations (even larger than the special iteration – in this research, typically 25 000 iterations) on the entire final population, and the best solution is found. Typically, a genetic algorithm would stop here and simply use this solution as its final answer. However, due to the random nature of the bootstrap, there is no guarantee that the final few generations did not actually degenerate and yield solutions inferior to those of earlier generations.

   As a result, we add the best solution of the final population to our collection of best portfolios, perform one final bootstrap with very large number of iterations on these portfolios, and select the best allocation out of this collection of portfolios as the final answer.

   Clearly these last two bootstrap operations will, of all the bootstrap operations, have the highest impact on our final solution. A larger number of bootstrap iterations is therefore warranted, as it will help distinguish more finely between a collection of already very competitive portfolios.

7.2. The optimal portfolios of the five asset case: more detailed results

7.2.1. Block length of one month (minimum standard error)

Table 38 and Table 39 show all the optimal portfolios for a block length of 1 month. As already discussed, these optimal portfolios minimise the standard errors of estimates of variance for SA equities, and are therefore of particular importance. For a brief discussion of these optimal portfolios, see Section 5.1.3.

criterion	horizon	req real return	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	long run return	annualised monthly stdev
MVO (req ret)		2%	2%	3%	90%	1%	4%	3%	5%	2.0%	1.6%
MVO (req ret)		4%	21%	9%	49%	1%	19%	23%	20%	4.0%	5.4%
MVO (req ret)		6%	41%	7%	10%	7%	35%	48%	42%	6.0%	10.4%
prob of success	12	2%	38%	13%	13%	10%	25%	48%	36%	5.7%	9.8%
prob of success	36	2%	45%	9%	3%	4%	39%	50%	43%	6.2%	11.2%
prob of success	60	2%	40%	8%	10%	8%	35%	48%	42%	5.9%	10.4%
prob of success	84	2%	39%	15%	4%	6%	37%	44%	43%	5.9%	10.2%
prob of success	120	2%	43%	5%	8%	4%	40%	47%	44%	6.1%	10.9%
prob of success	240	2%	40%	4%	17%	6%	33%	46%	39%	5.8%	10.1%
prob of success	12	4%	60%	1%	0%	22%	18%	82%	39%	7.3%	14.8%
prob of success	36	4%	63%	0%	0%	12%	25%	75%	37%	7.2%	14.6%
prob of success	60	4%	57%	0%	0%	15%	28%	72%	43%	7.1%	13.9%
prob of success	84	4%	58%	0%	0%	19%	23%	77%	42%	7.2%	14.3%
prob of success	120	4%	61%	0%	0%	7%	32%	68%	39%	7.1%	14.0%
prob of success	240	4%	57%	0%	0%	15%	28%	72%	43%	7.1%	13.9%
prob of success	12	6%	71%	0%	0%	28%	0%	100%	28%	7.6%	17.3%
prob of success	36	6%	74%	0%	0%	24%	1%	99%	26%	7.6%	17.6%
prob of success	60	6%	76%	0%	0%	23%	1%	99%	24%	7.7%	17.8%
prob of success	84	6%	71%	0%	0%	21%	8%	92%	29%	7.5%	16.7%
prob of success	120	6%	76%	0%	0%	24%	0%	100%	24%	7.7%	17.8%
prob of success	240	6%	74%	0%	0%	25%	1%	99%	26%	7.6%	17.6%
avg shortfall	12	2%	7%	0%	89%	1%	4%	8%	5%	2.3%	2.0%
avg shortfall	36	2%	11%	0%	79%	2%	8%	13%	10%	2.9%	3.0%
avg shortfall	60	2%	15%	0%	73%	2%	11%	16%	13%	3.2%	3.7%
avg shortfall	84	2%	17%	3%	66%	2%	13%	19%	14%	3.4%	4.2%
avg shortfall	120	2%	19%	1%	62%	3%	14%	22%	17%	3.7%	4.9%
avg shortfall	240	2%	25%	1%	50%	4%	20%	29%	24%	4.3%	6.3%
avg shortfall	12	4%	20%	11%	48%	4%	17%	24%	21%	4.0%	5.5%
avg shortfall	36	4%	22%	2%	55%	3%	17%	26%	21%	4.0%	5.6%
avg shortfall	60	4%	30%	5%	37%	4%	24%	34%	28%	4.8%	7.5%
avg shortfall	84	4%	36%	4%	27%	5%	29%	41%	34%	5.4%	8.9%
avg shortfall	120	4%	41%	9%	8%	7%	35%	47%	42%	6.0%	10.5%
avg shortfall	240	4%	52%	0%	0%	6%	42%	58%	48%	6.8%	12.7%
avg shortfall	12	6%	41%	3%	15%	10%	31%	50%	41%	6.0%	10.5%
avg shortfall	36	6%	42%	3%	13%	5%	36%	47%	41%	6.0%	10.5%
avg shortfall	60	6%	45%	8%	2%	6%	39%	51%	45%	6.3%	11.4%
avg shortfall	84	6%	51%	1%	0%	8%	40%	59%	48%	6.7%	12.6%
avg shortfall	120	6%	54%	0%	0%	9%	37%	63%	46%	6.9%	13.1%
avg shortfall	240	6%	58%	0%	0%	12%	30%	70%	42%	7.1%	13.8%
ds dev	12	2%	5%	0%	93%	1%	1%	6%	3%	2.1%	1.7%
ds dev	36	2%	8%	0%	86%	2%	4%	10%	6%	2.5%	2.3%
ds dev	60	2%	10%	0%	82%	2%	7%	12%	8%	2.7%	2.7%
ds dev	84	2%	12%	0%	79%	3%	7%	14%	10%	2.9%	3.1%
ds dev	120	2%	14%	0%	74%	2%	10%	16%	12%	3.1%	3.5%
ds dev	240	2%	18%	2%	63%	2%	15%	20%	16%	3.6%	4.6%
ds dev	12	4%	22%	5%	54%	1%	17%	24%	19%	4.0%	5.5%
ds dev	36	4%	20%	3%	54%	4%	19%	24%	22%	4.0%	5.4%
ds dev	60	4%	22%	3%	56%	3%	17%	25%	20%	4.0%	5.4%
ds dev	84	4%	22%	2%	56%	3%	17%	25%	20%	4.0%	5.5%
ds dev	120	4%	26%	5%	44%	3%	22%	29%	26%	4.4%	6.6%
ds dev	240	4%	38%	5%	20%	7%	31%	45%	37%	5.7%	9.7%
ds dev	12	6%	40%	6%	8%	6%	41%	45%	46%	6.0%	10.5%
ds dev	36	6%	41%	8%	9%	5%	37%	46%	42%	6.0%	10.5%
ds dev	60	6%	41%	11%	6%	6%	37%	46%	43%	6.0%	10.5%
ds dev	84	6%	41%	11%	5%	3%	40%	44%	43%	6.0%	10.5%
ds dev	120	6%	41%	7%	9%	7%	37%	48%	43%	6.0%	10.6%
ds dev	240	6%	52%	0%	0%	6%	41%	59%	47%	6.8%	12.8%

Table 38: Optimal portfolios with a block length of 1 month for all problem settings

criterion	horizon	req real return	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	long run return	annualised monthly stdev
max drawdown	12	2%	2%	0%	92%	0%	5%	2%	5%	2.0%	1.6%
max drawdown	36	2%	2%	0%	92%	0%	6%	2%	6%	2.0%	1.6%
max drawdown	60	2%	2%	1%	91%	1%	5%	3%	6%	2.0%	1.6%
max drawdown	84	2%	1%	0%	92%	2%	4%	4%	6%	2.0%	1.6%
max drawdown	120	2%	2%	0%	91%	0%	6%	2%	6%	2.0%	1.6%
max drawdown	240	2%	2%	0%	92%	0%	6%	2%	6%	2.0%	1.6%
max drawdown	12	4%	23%	6%	53%	0%	18%	23%	18%	4.0%	5.5%
max drawdown	36	4%	17%	16%	38%	8%	21%	25%	29%	4.0%	5.8%
max drawdown	60	4%	21%	5%	50%	0%	24%	21%	24%	4.0%	5.5%
max drawdown	84	4%	15%	19%	34%	7%	24%	22%	31%	4.0%	5.8%
max drawdown	120	4%	20%	3%	50%	0%	27%	20%	27%	4.0%	5.7%
max drawdown	240	4%	20%	2%	51%	0%	27%	20%	27%	4.0%	5.6%
max drawdown	12	6%	40%	11%	4%	5%	40%	45%	45%	6.0%	10.5%
max drawdown	36	6%	41%	5%	10%	5%	39%	46%	44%	6.0%	10.5%
max drawdown	60	6%	41%	0%	15%	5%	38%	46%	44%	6.0%	10.5%
max drawdown	84	6%	40%	8%	9%	7%	36%	47%	43%	6.0%	10.4%
max drawdown	120	6%	40%	11%	1%	2%	47%	42%	49%	6.0%	10.7%
max drawdown	240	6%	40%	3%	6%	1%	51%	41%	51%	6.0%	10.8%
sd horizon	12	2%	3%	0%	93%	0%	3%	3%	3%	2.0%	1.5%
sd horizon	36	2%	2%	2%	93%	3%	0%	5%	3%	2.0%	1.5%
sd horizon	60	2%	2%	0%	94%	3%	1%	5%	4%	2.0%	1.5%
sd horizon	84	2%	3%	0%	94%	0%	2%	4%	3%	2.0%	1.5%
sd horizon	120	2%	4%	1%	94%	0%	2%	4%	2%	2.0%	1.5%
sd horizon	240	2%	3%	1%	94%	2%	0%	5%	2%	2.0%	1.5%
sd horizon	12	4%	21%	19%	40%	0%	20%	21%	20%	4.0%	5.6%
sd horizon	36	4%	21%	12%	45%	0%	21%	21%	21%	4.0%	5.5%
sd horizon	60	4%	20%	21%	37%	0%	22%	20%	22%	4.0%	5.6%
sd horizon	84	4%	18%	23%	35%	3%	22%	21%	24%	4.0%	5.6%
sd horizon	120	4%	21%	9%	50%	4%	16%	25%	20%	4.0%	5.5%
sd horizon	240	4%	21%	18%	40%	0%	21%	21%	22%	4.0%	5.6%
sd horizon	12	6%	41%	10%	7%	5%	38%	46%	42%	6.0%	10.5%
sd horizon	36	6%	45%	9%	11%	0%	35%	45%	35%	6.0%	10.5%
sd horizon	60	6%	42%	0%	17%	7%	35%	48%	41%	6.0%	10.5%
sd horizon	84	6%	40%	8%	9%	7%	35%	48%	43%	6.0%	10.5%
sd horizon	120	6%	41%	3%	14%	7%	34%	48%	41%	6.0%	10.5%
sd horizon	240	6%	39%	8%	7%	7%	39%	46%	46%	6.0%	10.5%

Table 39: Optimal portfolios with a block length of 1 month for all problem settings (continued)

7.2.2. Empirical results (block length fixed to investment horizon)

On the other extreme, this section considers the longest possible block length. The term "empirical case" is used to denote the case where the block length is deterministically set to be equal to the horizon. In other words, the optimal portfolio in this scenario denotes the static allocation that would have actually fared the best over chronological, historical data for various investment horizons and risk criteria.

In this setting there is no simulation taking place, history speaks for itself. Essentially all the "rolling periods" of investment horizon length are examined, and the portfolio that performed the best according to the chosen risk metric, with at least a real return of 2%, 4% or 6%, is found by the genetic algorithm. Hence, these results are very interesting in and of themselves as a documentation of history, as they are the historically optimal portfolios for each problem setting. These portfolios are tabulated in Table 40 and Table 41 for ease of reference to the interested reader.

As discussed in Section 4.3.4, the choice of block length is, generally speaking, a bias-variance trade-off in the estimation of the underlying joint distribution of the asset classes. If the block length is set deterministically to the length of the investment horizon, the estimated joint distribution should be unbiased, though at the cost of higher variance.

criterion	horizon	req real return	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	long run return	annualised monthly stdev	% improv on MVO (avg return)	% improv on MVO (req return)
MVO (req return)		2%	2%	3%	90%	1%	4%	3%	5%	2.0%	1.6%		
MVO (req return)		4%	21%	9%	49%	1%	19%	23%	20%	4.0%	5.4%		
MVO (req return)		6%	41%	7%	10%	7%	35%	48%	42%	6.0%	10.4%		
avg shortfall	12	2%	16%	0%	78%	0%	6%	16%	6%	3.1%	3.6%	4.6%	18%
avg shortfall	36	2%	33%	0%	52%	0%	15%	33%	15%	4.5%	7.1%	12.6%	57%
avg shortfall	60	2%	40%	0%	44%	0%	16%	40%	16%	5.1%	8.6%	17.8%	78%
avg shortfall	84	2%	68%	1%	26%	1%	4%	69%	4%	6.4%	14.4%	87.7%	99%
avg shortfall	120	2%	95%	2%	0%	2%	1%	96%	3%	7.7%	20.1%	100.0%	100%
avg shortfall	240	2%	23%	16%	18%	21%	22%	44%	43%	5.1%	8.5%	0.0%	100%
avg shortfall	36	4%	39%	0%	42%	0%	19%	39%	19%	5.1%	8.5%	10.1%	19%
avg shortfall	60	4%	55%	0%	16%	0%	29%	55%	29%	6.4%	12.0%	12.0%	47%
avg shortfall	12	4%	25%	0%	61%	0%	14%	25%	14%	4.0%	5.5%	3.3%	3%
avg shortfall	84	4%	73%	15%	0%	9%	2%	83%	12%	7.1%	16.4%	46.5%	84%
avg shortfall	120	4%	83%	0%	0%	0%	17%	83%	17%	7.6%	17.4%	34.7%	95%
avg shortfall	240	4%	32%	3%	10%	37%	17%	69%	54%	6.2%	11.8%	0.0%	100%
avg shortfall	36	6%	48%	2%	19%	0%	31%	48%	31%	6.0%	10.7%	5.4%	6%
avg shortfall	60	6%	62%	4%	0%	0%	34%	62%	34%	6.9%	13.6%	12.9%	20%
avg shortfall	84	6%	76%	3%	0%	0%	20%	76%	20%	7.3%	16.2%	14.4%	43%
avg shortfall	120	6%	72%	0%	0%	12%	16%	84%	28%	7.5%	16.1%	13.5%	52%
avg shortfall	12	6%	43%	10%	6%	0%	41%	43%	41%	6.0%	10.6%	0.9%	0%
avg shortfall	240	6%	76%	0%	0%	10%	13%	87%	24%	7.5%	16.9%	51.0%	94%
ds dev	12	2%	10%	0%	87%	2%	1%	12%	3%	2.6%	2.5%	2.1%	14%
ds dev	36	2%	28%	0%	49%	0%	23%	28%	23%	4.4%	6.6%	4.1%	43%
ds dev	60	2%	34%	0%	49%	0%	16%	34%	16%	4.7%	7.5%	8.7%	59%
ds dev	84	2%	71%	0%	25%	0%	4%	71%	4%	6.6%	14.8%	72.1%	95%
ds dev	120	2%	91%	1%	3%	0%	5%	91%	5%	7.5%	19.1%	100.0%	100%
ds dev	240	2%	31%	22%	23%	17%	7%	48%	24%	5.1%	8.7%	0.0%	100%
ds dev	36	4%	31%	0%	48%	0%	21%	31%	21%	4.6%	7.1%	4.4%	11%
ds dev	60	4%	43%	0%	37%	0%	20%	43%	20%	5.4%	9.4%	7.5%	27%
ds dev	12	4%	23%	3%	55%	1%	19%	23%	19%	4.0%	5.5%	1.3%	1%
ds dev	84	4%	75%	12%	0%	5%	8%	79%	13%	7.1%	16.3%	35.1%	74%
ds dev	120	4%	88%	0%	0%	0%	12%	88%	12%	7.6%	18.4%	28.4%	86%
ds dev	240	4%	39%	1%	3%	30%	27%	69%	57%	6.6%	12.6%	0.0%	100%
ds dev	36	6%	45%	7%	12%	0%	36%	45%	36%	6.0%	10.6%	4.5%	4%
ds dev	60	6%	52%	0%	19%	0%	28%	52%	28%	6.2%	11.4%	6.6%	8%
ds dev	84	6%	76%	11%	0%	3%	10%	79%	14%	7.2%	16.3%	17.8%	41%
ds dev	120	6%	75%	0%	0%	11%	14%	86%	25%	7.5%	16.6%	9.3%	45%
ds dev	12	6%	44%	4%	13%	0%	39%	44%	39%	6.0%	10.6%	2.7%	1%
ds dev	240	6%	76%	0%	0%	13%	11%	89%	24%	7.6%	16.9%	41.7%	88%
prob of success	12	2%	44%	4%	1%	22%	28%	66%	50%	6.6%	12.5%	1.3%	40%
prob of success	36	2%	42%	25%	13%	0%	20%	42%	20%	5.4%	9.6%	10.3%	52%
prob of success	60	2%	48%	26%	7%	2%	17%	49%	19%	5.8%	10.9%	3.6%	68%
prob of success	84	2%	68%	6%	16%	7%	3%	75%	10%	6.7%	14.9%	5.9%	79%
prob of success	120	2%	92%	0%	6%	1%	0%	94%	2%	7.5%	19.5%	0.7%	59%
prob of success	240	2%	11%	21%	4%	44%	20%	55%	65%	5.1%	10.5%	0.0%	32%
prob of success	36	4%	58%	0%	0%	41%	0%	99%	41%	7.4%	16.2%	3.5%	33%
prob of success	60	4%	61%	5%	1%	0%	33%	62%	33%	6.9%	13.5%	5.1%	30%
prob of success	12	4%	52%	3%	3%	17%	26%	68%	42%	6.8%	13.0%	1.5%	29%
prob of success	84	4%	79%	12%	0%	1%	8%	80%	9%	7.2%	16.9%	4.6%	39%
prob of success	120	4%	65%	1%	0%	10%	25%	75%	35%	7.2%	14.7%	0.8%	37%
prob of success	240	4%	35%	6%	2%	41%	17%	76%	58%	6.5%	12.8%	0.0%	21%
prob of success	36	6%	53%	0%	0%	47%	0%	100%	47%	7.3%	15.9%	6.1%	18%
prob of success	60	6%	64%	7%	1%	2%	25%	67%	28%	7.0%	14.1%	11.1%	18%
prob of success	84	6%	74%	2%	1%	6%	18%	80%	24%	7.4%	16.1%	3.2%	17%
prob of success	120	6%	55%	4%	0%	4%	37%	59%	41%	6.8%	12.9%	5.0%	11%
prob of success	12	6%	60%	1%	2%	35%	3%	95%	38%	7.3%	15.8%	1.0%	12%
prob of success	240	6%	67%	0%	0%	8%	25%	75%	33%	7.3%	15.0%	1.4%	11%

Table 40: The empirically optimal portfolios (block length=horizon length)

criterion	horizon	req real return	SA equities	SA bonds	SA cash	Global equities	Global bonds	Total equities	Total offshore	long run return	annualised monthly stdev	% improv on MVO (avg return)	% improv on MVO (req return)
sd horison	12	2%	7%	14%	74%	0%	5%	7%	5%	2.4%	2.4%	4.1%	5%
sd horison	36	2%	15%	38%	42%	0%	6%	15%	6%	3.1%	4.9%	13.0%	13%
sd horison	60	2%	26%	62%	0%	0%	12%	26%	12%	4.3%	8.3%	26.4%	21%
sd horison	84	2%	27%	72%	0%	0%	0%	27%	0%	4.0%	9.2%	29.5%	24%
sd horison	120	2%	34%	66%	0%	0%	0%	34%	0%	4.4%	10.0%	41.4%	30%
sd horison	240	2%	5%	87%	0%	0%	8%	5%	8%	2.6%	7.3%	42.1%	38%
sd horison	36	4%	22%	48%	17%	0%	14%	22%	14%	4.0%	6.8%	15.9%	9%
sd horison	60	4%	26%	62%	0%	0%	12%	26%	12%	4.3%	8.3%	26.4%	18%
sd horison	12	4%	21%	31%	30%	0%	17%	21%	17%	4.0%	6.0%	8.7%	3%
sd horison	84	4%	27%	72%	0%	0%	0%	27%	0%	4.0%	9.2%	29.5%	23%
sd horison	120	4%	34%	66%	0%	0%	0%	34%	0%	4.4%	10.0%	41.4%	35%
sd horison	240	4%	27%	73%	0%	0%	0%	27%	0%	4.0%	9.2%	48.2%	42%
sd horison	36	6%	47%	26%	0%	0%	26%	48%	27%	6.0%	11.0%	12.9%	14%
sd horison	60	6%	50%	31%	0%	0%	19%	50%	19%	6.0%	11.5%	20.8%	16%
sd horison	84	6%	54%	33%	0%	0%	13%	54%	13%	6.0%	12.3%	22.0%	22%
sd horison	120	6%	61%	39%	0%	0%	0%	61%	0%	6.0%	13.9%	38.1%	34%
sd horison	12	6%	47%	10%	12%	0%	30%	47%	30%	6.0%	10.6%	6.3%	5%
sd horison	240	6%	60%	40%	0%	0%	0%	60%	0%	6.0%	13.8%	42.5%	42%
max drawdown	12	2%	2%	0%	92%	0%	5%	2%	6%	2.0%	1.6%	94.8%	88%
max drawdown	36	2%	3%	0%	93%	0%	4%	3%	4%	2.0%	1.5%	77.2%	88%
max drawdown	60	2%	3%	0%	92%	0%	5%	3%	5%	2.0%	1.6%	78.9%	55%
max drawdown	84	2%	2%	0%	92%	0%	5%	3%	5%	2.0%	1.6%	85.2%	94%
max drawdown	120	2%	2%	0%	92%	0%	5%	3%	6%	2.0%	1.6%	74.7%	42%
max drawdown	240	2%	2%	4%	90%	0%	4%	3%	4%	2.0%	1.5%	77.6%	51%
max drawdown	36	4%	20%	7%	48%	1%	24%	21%	25%	4.0%	5.5%	5.3%	4%
max drawdown	60	4%	19%	13%	39%	0%	29%	19%	29%	4.0%	5.7%	2.2%	1%
max drawdown	12	4%	19%	11%	48%	6%	16%	25%	23%	4.0%	5.5%	-1.5%	-2%
max drawdown	84	4%	18%	13%	42%	4%	22%	22%	26%	4.0%	5.5%	4.8%	6%
max drawdown	120	4%	17%	17%	39%	5%	23%	22%	28%	4.0%	5.6%	3.8%	8%
max drawdown	240	4%	15%	26%	29%	5%	25%	20%	30%	4.0%	5.8%	8.2%	17%
max drawdown	36	6%	42%	2%	12%	1%	43%	43%	44%	6.0%	10.5%	8.1%	8%
max drawdown	60	6%	44%	13%	7%	0%	36%	44%	36%	6.0%	10.5%	6.2%	3%
max drawdown	84	6%	40%	5%	5%	0%	49%	40%	49%	6.0%	10.7%	10.9%	6%
max drawdown	120	6%	41%	3%	9%	0%	48%	41%	48%	6.0%	10.6%	9.0%	0%
max drawdown	12	6%	43%	5%	11%	0%	41%	43%	41%	6.0%	10.5%	3.6%	2%
max drawdown	240	6%	44%	16%	2%	0%	37%	44%	37%	6.0%	10.6%	10.6%	7%

Table 41: The empirically optimal portfolios (block length=horizon length) (continued).

For average shortfall, downside deviation, and standard deviation over horizon, with a few exceptions, there is a clear general pattern in Table 40 and Table 41 of a longer investment horizon resulting in a higher allocation to equities and a more volatile optimal portfolio in terms of monthly standard deviation. One repeated exception to this pattern is the 240-month investment horizon, where the optimal portfolio often drops off in terms of equity allocation (though it still significantly higher than those of empirically optimal portfolios for 12 and 36 month horizons). Consider, however, that for a dataset of 541 months, there are a limited number of rolling periods when the horizon is 240 months:

block length	Number of rollling periods	mutually exclusive rolling periods
12	530	44.2
36	506	14.1
60	482	8.0
84	458	5.5
120	422	3.5
240	302	1.3

Table 42: The number of definable units for increasing block lengths

The empirical portfolio for a horizon of 240 months is based on only 1.3 mutually exclusive periods in the total dataset and this small sample size is the most likely explanation of that seemingly anomalous optimal portfolio.

It is important to note that the optimal portfolios performed far better than the MVO portfolios (with the required real return) with respect to the risk metric that they minimise. The same can be said for the probability of success portfolio. This is in contrast to the maximum drawdown optimal portfolio, which perhaps does not materially reduce the average maximum drawdown. For example, see Table 43, an extract showing only the required real return of 4% and horizon of 84 months scenario:

criterion	horizon	req real return	SA equities	SA bonds	SA cash	Global equities	Global bonds	risk criterion optimised	risk criterion of MVO (avg return)	risk criterion of MVO (req return)	% improv on MVO (avg return)	% improv on MVO (req return)
MVO (req return)	84	4%	21%	9%	49%	1%	19%	5.40%				
avg shortfall	84	4%	73%	15%	0%	9%	2%	0.02%	0.03%	0.11%	47%	84%
ds dev	84	4%	75%	12%	0%	5%	8%	0.07%	0.10%	0.25%	35%	74%
max drawdown	84	4%	18%	13%	42%	4%	22%	4.71%	4.95%	5.00%	5%	6%
prob of success	84	4%	79%	12%	0%	1%	8%	89.31%	85.37%	64.41%	5%	39%
sd horison	84	4%	27%	72%	0%	0%	0%	0.28%	0.40%	0.36%	30%	23%

Table 43: Excerpt from empirical case (required real return of 4% and horizon of 84 months)

The last two columns of Table 43 show the percentage improvements gained by veering from the MVO portfolios. Compared with the MVO portfolio (that meets the required return), the average shortfall portfolio reduces its target metric from 0.11% to 0.02%, the downside deviation portfolio similarly from 0.25% to 0.07%, the probability of success increases from 64% to 89% – all of these improvements seem material. On the other hand, the maximum drawdown improvement from -5.0% to only -4.7%, hardly seems important from a practical standpoint.

In other words, the results indicate that if an investor is averse to the average shortfall, downside deviation, or probability of success over investment horizon, history suggests there may be something for her to gain by not simply adhering to the MVO portfolio that meets the required return. Consistent with the results for intermediate block length, it appears doubtful optimising for maximum drawdown yields any real and exploitable benefits – it would appear that empirically over our dataset the minimum standard deviation portfolio is a good approximation of the minimum drawdown portfolio.

Lastly we note that, consistent with the surprising result seen for the intermediate block length, the optimal probability of success portfolio follows no clear pattern with respect to investment horizon. What is clear, however, is that the optimal portfolio for this metric was far more aggressive than the MVO portfolio, and also increased the probability of success materially, for all horizons longer than 12 months.

7.3. The optimal portfolios of the eight asset case (ignoring regimes)

This section briefly considers the case where we include SA property, global property and global cash, to get a sense of the role these asset classes could play in a strategic sense. In other words, we are now optimising to the following asset classes: SA equities, SA bonds, SA cash, SA property, global equities, global bonds, global cash and global property.

For this set of optimisations, we solved the following settings:

- Real return of 2%,4%, and 6%;
- Block length that is defined as a ratio of the investment horizon: 0% (rounded up to 1 month), 25%,75%, and 100% (empirical);
- Investment horizon of 12, 36, 60, 84, 120, 240, and 360 months; and
- All the risk criteria discussed in our main results.

Due to the fact that we considered only block lengths with a specific ratio of the horizon, we cannot choose exactly the same block lengths as before as our intermediate block length. However, we chose similar lengths for each horizon among those available, shown in Table 44:

horizon	optimal block length
12	9
36	18
60	30
84	42
120	30
240	60

Table 44: Intermediate block lengths for eight asset case

As the choices are broadly similar to those of the five asset case, we will not motivate them again in any detail here. Refer back to Section 4.3.4 for the discussion on this topic.

7.3.1. The eight asset case: Results for the intermediate block length

In the Table 45 below we show the optimal portfolios of the intermediate block lengths and for investment horizons of 12 months, 60 months, 120 months and 240 months.

criterion	horizon	required real return	SA equities	SA bonds	SA cash	SA property	Global equities	Global bonds	Global cash	Global property	total equities	total Global	total property	real return	annualised monthly stdev
MVO req return			2%	0%	94%	1%	0%	1%	0%	1.0%	2%	2%	2%	2.0%	1.4%
MVO req return			13%	3%	54%	11%	0%	16%	0%	3.2%	13%	20%	14%	4.0%	4.9%
MVO req return			21%	9%	10%	23%	2%	24%	0%	12.4%	23%	38%	35%	6.0%	9.3%
avg shortfall	12	2.0%	11%	0%	80%	2%	0%	7%	0%	0.0%	11%	7%	2%	2.8%	2.9%
avg shortfall	60	2.0%	33%	0%	49%	0%	0%	18%	0%	0.0%	33%	18%	0%	4.7%	7.3%
avg shortfall	120	2.0%	43%	0%	32%	0%	0%	25%	0%	0.0%	43%	25%	0%	5.5%	9.5%
avg shortfall	240	2.0%	59%	0%	6%	0%	0%	35%	0%	0.0%	59%	35%	0%	6.8%	13.0%
avg shortfall	12	4.0%	20%	0%	57%	5%	0%	16%	0%	0.1%	21%	17%	5%	4.0%	5.3%
avg shortfall	60	4.0%	45%	0%	28%	0%	0%	27%	0%	0.0%	45%	27%	0%	5.7%	10.0%
avg shortfall	120	4.0%	60%	0%	1%	0%	0%	39%	0%	0.0%	60%	39%	0%	6.9%	13.5%
avg shortfall	240	4.0%	68%	0%	0%	0%	0%	32%	0%	0.0%	68%	32%	0%	7.2%	14.7%
avg shortfall	12	6.0%	30%	0%	12%	18%	0%	39%	0%	0.0%	30%	39%	18%	6.1%	9.7%
avg shortfall	60	6.0%	56%	1%	2%	2%	0%	39%	0%	0.0%	56%	39%	2%	6.8%	12.9%
avg shortfall	120	6.0%	65%	0%	0%	0%	0%	35%	0%	0.0%	65%	35%	0%	7.1%	14.2%
avg shortfall	240	6.0%	74%	0%	0%	0%	1%	25%	0%	0.0%	75%	26%	0%	7.4%	15.8%
ds dev	12	2.0%	8%	0%	88%	1%	1%	1%	0%	0.3%	9%	3%	2%	2.5%	2.3%
ds dev	60	2.0%	26%	0%	60%	0%	0%	14%	0%	0.0%	26%	14%	0%	4.1%	5.8%
ds dev	120	2.0%	35%	0%	44%	0%	0%	20%	0%	0.0%	35%	20%	0%	4.9%	7.8%
ds dev	240	2.0%	53%	0%	20%	0%	0%	26%	0%	0.0%	53%	26%	0%	6.2%	11.6%
ds dev	12	4.0%	17%	1%	54%	7%	0%	22%	0%	0.0%	17%	22%	7%	4.0%	5.2%
ds dev	60	4.0%	33%	0%	48%	0%	0%	19%	0%	0.0%	33%	19%	0%	4.7%	7.3%
ds dev	120	4.0%	47%	0%	25%	0%	0%	28%	0%	0.0%	47%	28%	0%	5.8%	10.3%
ds dev	240	4.0%	65%	0%	0%	0%	0%	35%	0%	0.0%	65%	35%	0%	7.1%	14.2%
ds dev	12	6.0%	27%	1%	12%	20%	0%	37%	0%	2.7%	27%	40%	22%	6.0%	9.5%
ds dev	60	6.0%	48%	0%	20%	0%	0%	32%	0%	0.0%	48%	32%	0%	6.0%	10.8%
ds dev	120	6.0%	59%	0%	2%	0%	0%	39%	0%	0.0%	59%	39%	0%	6.9%	13.1%
ds dev	240	6.0%	70%	0%	0%	0%	0%	30%	0%	0.0%	70%	30%	0%	7.3%	15.0%
prob of suc	12	2.0%	29%	0%	6%	15%	25%	11%	0%	12.2%	55%	49%	28%	6.7%	12.0%
prob of suc	60	2.0%	51%	5%	9%	5%	0%	31%	0%	0.0%	51%	31%	5%	6.4%	11.7%
prob of suc	120	2.0%	58%	1%	3%	0%	0%	37%	0%	0.0%	59%	37%	0%	6.8%	13.0%
prob of suc	240	2.0%	65%	0%	1%	0%	0%	33%	0%	0.2%	65%	34%	0%	7.1%	14.2%
prob of suc	12	4.0%	22%	0%	0%	14%	26%	0%	0%	38.4%	48%	64%	52%	6.9%	13.7%
prob of suc	60	4.0%	58%	0%	0%	13%	0%	28%	0%	0.1%	59%	28%	13%	7.3%	13.9%
prob of suc	120	4.0%	66%	0%	0%	2%	1%	30%	0%	0.1%	67%	31%	2%	7.2%	14.5%
prob of suc	240	4.0%	70%	0%	0%	0%	1%	28%	0%	0.0%	71%	30%	0%	7.3%	15.1%
prob of suc	12	6.0%	34%	0%	0%	10%	24%	0%	0%	31.5%	59%	56%	41%	7.2%	14.0%
prob of suc	60	6.0%	45%	0%	0%	25%	19%	2%	0%	8.7%	64%	29%	34%	7.5%	14.6%
prob of suc	120	6.0%	70%	0%	0%	6%	16%	2%	0%	5.0%	87%	23%	11%	7.7%	16.9%
prob of suc	240	6.0%	79%	0%	0%	1%	12%	4%	0%	4.0%	91%	20%	5%	7.7%	17.8%
sd horizon	12	2.0%	3%	2%	92%	3%	0%	0%	1%	0.0%	3%	1%	3%	2.0%	1.6%
sd horizon	60	2.0%	6%	10%	75%	5%	0%	0%	4%	0.0%	6%	4%	5%	2.4%	2.5%
sd horizon	120	2.0%	5%	11%	72%	6%	0%	0%	6%	0.0%	5%	6%	6%	2.3%	2.6%
sd horizon	240	2.0%	8%	17%	51%	9%	0%	0%	15%	0.0%	8%	15%	9%	2.5%	3.9%
sd horizon	12	4.0%	15%	6%	53%	10%	1%	14%	0%	0.2%	17%	15%	10%	4.0%	5.1%
sd horizon	60	4.0%	25%	20%	46%	1%	0%	5%	0%	2.0%	25%	8%	3%	4.0%	6.2%
sd horizon	120	4.0%	24%	16%	52%	6%	2%	1%	0%	0.0%	25%	2%	6%	4.0%	6.2%
sd horizon	240	4.0%	24%	24%	45%	6%	0%	0%	0%	0.2%	25%	1%	6%	4.0%	6.7%
sd horizon	12	6.0%	32%	3%	17%	17%	0%	29%	0%	2.3%	32%	32%	19%	6.0%	9.4%
sd horizon	60	6.0%	41%	20%	4%	7%	0%	27%	0%	0.2%	41%	28%	7%	6.0%	10.3%
sd horizon	120	6.0%	49%	24%	5%	0%	0%	20%	0%	1.9%	49%	23%	2%	6.0%	11.2%
sd horizon	240	6.0%	47%	26%	6%	8%	0%	10%	1%	2.0%	47%	13%	10%	6.0%	11.6%
max drawdown	12	2.0%	2%	1%	91%	1%	0%	5%	0%	0.0%	2%	6%	1%	2.0%	1.6%
max drawdown	60	2.0%	0%	2%	90%	3%	1%	5%	0%	0.0%	1%	6%	3%	2.0%	1.6%
max drawdown	120	2.0%	1%	1%	90%	2%	0%	5%	0%	0.2%	1%	6%	2%	2.0%	1.6%
max drawdown	240	2.0%	0%	1%	91%	3%	0%	5%	0%	0.1%	0%	5%	3%	2.0%	1.6%
max drawdown	12	4.0%	13%	1%	54%	11%	0%	21%	0%	0.2%	13%	21%	12%	4.0%	5.0%
max drawdown	60	4.0%	14%	2%	54%	11%	0%	17%	0%	1.8%	14%	19%	13%	4.0%	5.0%
max drawdown	120	4.0%	13%	3%	51%	11%	0%	21%	0%	0.0%	13%	21%	11%	4.0%	5.0%
max drawdown	240	4.0%	16%	13%	41%	5%	0%	25%	0%	0.5%	16%	25%	5%	4.0%	5.3%
max drawdown	12	6.0%	28%	2%	13%	22%	1%	33%	0%	1.3%	29%	36%	23%	6.1%	9.5%
max drawdown	60	6.0%	26%	1%	16%	21%	0%	32%	0%	2.5%	27%	35%	24%	5.9%	9.2%
max drawdown	120	6.0%	25%	4%	9%	22%	0%	38%	0%	1.5%	25%	40%	24%	6.0%	9.4%
max drawdown	240	6.0%	27%	6%	8%	20%	0%	40%	0%	0.0%	27%	40%	20%	6.0%	9.4%

Table 45: All optimal portfolios for 8 asset case (intermediate block lengths)

Global cash

The global cash weight is consistently very small in Table 45. The average allocation across all of the optimal portfolios is a mere 1.2%. That is not to say by any means that the asset class does not have any role to play in *tactical* asset allocation – there would almost surely have been instances where holding this asset class would have improved risk-adjusted returns.

SA and global property

We have already motivated our reasons for excluding these asset classes from our main results: the fact that they are similar to equities in their return and risk characteristics and the two related issues of the small size of the asset class and lack of liquidity.

Typically, historical analysis shows that listed property plays an important part in optimal portfolios. The optimal allocations of 11% and 23% seen in the MVO portfolios (that are based on standard deviations of monthly returns) in the results above for required real returns of 4% and 6% are broadly representative of typical results. However, such high allocations are often met with scepticism, due to the small size and limited liquidity of this asset class.

However, it would appear that historically, the benefits of holding this asset class were diminished for longer investment horizons (than the one month implied by the MVO portfolio). Table 46 below compares the average allocation to SA property in the optimal portfolios for the four horizon sensitive criteria, namely average shortfall, downside deviation, probability of success and standard deviation over horizon, to that of MVO for the different required returns:

req return	4 criteria	MVO
2%	3%	1%
4%	4%	11%
6%	7%	23%
average	5%	12%

Table 46: Optimal SA property weight (MVO & other criteria)

In other words, the average optimal allocation reduces from 12% to 5% when investment horizon and longer periodicities are accounted for. A similar result can be seen for global property in Table 47:

req return	4 criteria	MVO
2%	1%	1%
4%	3%	3%
6%	12%	4%
average	6%	2%

Table 47: Optimal global property weight (MVO & other criteria)

The reason for the smaller role played by these asset classes over the longer investment horizon is likely the fact that SA property is increasingly dominated by SA equities over a longer and longer periodicity. As we saw in Section 4.1.4.2 the mean reversion seen in SA equities over longer periodicities makes it relatively less variable over the longer term. SA property is also highly correlated to SA equities, and thus offers little in the way of diversification benefits in a portfolio dominated by SA equities.

7.4. Regime-cognisant optimal portfolios of three risk criteria and various block lengths

Optimal regime-cognisant portfolios for a required real return of 6% and investment horizon of 36 months, for three different block lengths (one, 12, 24 months), backtested with returns data from 1972 to 2017 (see Section 5.2.3 for a discussion of these portfolios).

	monthly standard deviation											
	block length = 1 month				block length = 12 months				block length = 24 months			
	Update AA every 6 months											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
return p.a.	14.0%	12.5%	12.5%	13.3%	13.8%	13.0%	13.3%	13.4%	13.9%	13.4%	13.7%	13.7%
real return p.a.	5.8%	4.4%	4.3%	5.1%	5.6%	4.8%	5.2%	5.3%	5.8%	5.3%	5.5%	5.6%
avg shortfall	2.0%	2.2%	2.1%	1.6%	1.9%	1.7%	1.5%	1.6%	1.9%	1.6%	1.4%	1.4%
stdev of 36-month returns	4.9%	4.5%	4.2%	3.9%	4.8%	3.8%	3.8%	4.0%	5.0%	3.9%	3.7%	4.0%
monthly stdev	8.9%	4.0%	4.3%	5.5%	8.9%	4.5%	5.2%	5.5%	8.8%	5.5%	6.3%	6.4%
max drawdown	20.8%	5.1%	5.4%	9.4%	20.0%	8.6%	10.7%	12.4%	19.9%	12.8%	16.5%	17.8%
Shapre ratio (36-month stdev)	61.7%	36.2%	37.9%	61.5%	60.6%	54.7%	62.8%	62.9%	60.3%	65.0%	74.5%	69.8%
Sharpe ratio (monthly stdev)	34.3%	40.8%	37.1%	43.6%	32.4%	45.7%	46.5%	45.4%	34.1%	45.9%	43.7%	43.9%
	Maintain AA for 36-months											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
avg return. p.a.	13.1%	11.5%	11.9%	12.3%	13.0%	11.9%	12.1%	12.3%	12.9%	12.2%	12.6%	12.6%
avg real return p.a.	5.4%	3.7%	4.2%	4.6%	5.3%	4.2%	4.4%	4.5%	5.2%	4.5%	4.9%	4.9%
avg shortfall	2.1%	2.9%	2.4%	2.2%	2.2%	2.4%	2.3%	2.2%	2.3%	2.3%	2.1%	2.0%
stdev of 36-month returns	4.2%	3.8%	3.4%	3.3%	4.2%	3.4%	3.3%	3.3%	4.2%	3.4%	3.4%	3.5%
avg monthly stdev	8.7%	3.7%	4.1%	5.3%	8.7%	4.4%	5.1%	5.5%	8.7%	5.4%	6.2%	6.3%
avg max drawdown	12.0%	6.1%	6.0%	7.7%	12.0%	6.6%	7.4%	8.0%	12.1%	8.1%	9.1%	9.3%
Shapre ratio (36-month stdev)	51.3%	14.2%	30.4%	42.4%	49.9%	30.9%	37.9%	41.5%	48.2%	37.8%	49.8%	48.8%
Sharpe ratio (monthly stdev)	25.1%	14.8%	25.5%	26.2%	23.9%	23.6%	24.5%	25.0%	23.2%	23.8%	27.0%	26.6%
	Average asset allocation											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
SA equities	39%	3%	10%	15%	40%	13%	16%	19%	40%	18%	23%	23%
SA bonds	9%	27%	28%	40%	9%	29%	33%	35%	9%	35%	37%	37%
SA cash	12%	43%	37%	17%	12%	35%	24%	20%	12%	22%	14%	13%
Global equities	4%	12%	8%	4%	4%	6%	4%	2%	3%	4%	2%	1%
Global bonds	35%	14%	18%	24%	35%	17%	23%	24%	36%	21%	25%	25%
Total equities	44%	15%	18%	19%	43%	19%	20%	21%	43%	22%	24%	25%
Total offshore	39%	26%	26%	28%	39%	23%	27%	27%	39%	24%	27%	27%

Table 48: Monthly standard deviation optimal portfolios for a required real return of 6% and investment horizon of 36 months, for three different block lengths (one, 12, 24 months), backtested with returns data from 1972 to 2017.

	average shortfall											
	block length = 1 month				block length = 12 months				block length = 24 months			
	Update AA every 6 months											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
return p.a.	13.9%	13.2%	13.6%	14.0%	13.8%	14.1%	14.1%	14.2%	13.8%	14.5%	14.4%	14.4%
real return p.a.	5.7%	5.1%	5.5%	5.8%	5.7%	6.0%	5.9%	6.1%	5.7%	6.3%	6.3%	6.3%
avg shortfall	1.8%	3.2%	2.1%	1.5%	1.7%	1.7%	1.5%	1.5%	1.8%	1.5%	1.4%	1.4%
stdev of 36-month returns	4.7%	7.8%	5.0%	4.1%	4.6%	4.6%	4.3%	4.4%	4.7%	4.5%	4.4%	4.3%
monthly stdev	8.8%	8.5%	7.4%	7.2%	8.9%	8.3%	7.8%	8.2%	9.0%	8.9%	8.5%	8.6%
max drawdown	20.3%	23.0%	16.8%	12.4%	20.5%	15.6%	14.9%	15.7%	20.3%	18.4%	17.3%	18.2%
Shapre ratio (36-month stdev)	63.8%	30.1%	54.5%	74.8%	64.1%	70.9%	74.2%	76.5%	62.8%	78.8%	80.4%	82.0%
Sharpe ratio (monthly stdev)	33.8%	27.3%	36.9%	42.5%	32.9%	38.7%	40.5%	40.8%	32.5%	40.1%	41.5%	41.2%
	Maintain AA for 36-months											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
avg return. p.a.	13.0%	11.6%	12.6%	12.9%	13.0%	13.0%	13.0%	13.1%	13.0%	13.4%	13.4%	13.4%
avg real return p.a.	5.3%	3.7%	4.9%	5.2%	5.3%	5.3%	5.3%	5.5%	5.3%	5.7%	5.7%	5.7%
avg shortfall	2.1%	4.1%	2.7%	2.1%	2.1%	2.3%	2.2%	2.1%	2.1%	2.0%	1.9%	1.9%
stdev of 36-month returns	4.1%	6.5%	4.3%	3.5%	4.1%	4.0%	3.7%	3.8%	4.1%	4.0%	3.8%	3.8%
avg monthly stdev	8.6%	8.0%	7.2%	7.1%	8.7%	8.2%	7.7%	8.0%	8.9%	8.7%	8.3%	8.4%
avg max drawdown	11.8%	13.3%	10.7%	10.2%	11.9%	11.4%	10.8%	11.3%	12.1%	12.2%	11.5%	11.7%
Shapre ratio (36-month stdev)	50.8%	9.9%	39.0%	56.8%	51.5%	52.2%	56.2%	59.1%	51.1%	61.8%	64.7%	64.5%
Sharpe ratio (monthly stdev)	24.2%	8.0%	23.5%	28.3%	23.8%	25.7%	27.3%	28.0%	23.7%	28.4%	29.7%	29.5%
	Average asset allocation											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
SA equities	41%	8%	20%	28%	44%	38%	34%	37%	45%	44%	41%	42%
SA bonds	5%	36%	34%	41%	2%	25%	30%	30%	1%	26%	28%	27%
SA cash	16%	0%	0%	0%	22%	0%	0%	1%	23%	0%	3%	4%
Global equities	2%	29%	13%	1%	0%	0%	0%	0%	0%	0%	0%	0%
Global bonds	36%	27%	32%	31%	33%	37%	36%	33%	30%	30%	27%	27%
Total equities	42%	37%	34%	29%	44%	38%	34%	37%	45%	44%	41%	42%
Total offshore	38%	56%	46%	32%	33%	37%	36%	33%	30%	30%	27%	27%

Table 49: Average shortfall optimal portfolios for a required real return of 6% and investment horizon of 36 months, for three different block lengths (one, 12, 24 months), backtested with returns data from 1972 to 2017.

	standard deviation over horizon											
	block length = 1 month				block length = 12 months				block length = 24 months			
	Update AA every 6 months											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
return p.a.	14.0%	12.5%	12.5%	13.3%	13.9%	12.9%	13.2%	13.1%	14.2%	13.2%	13.6%	13.7%
real return p.a.	5.8%	4.4%	4.3%	5.1%	5.7%	4.8%	5.0%	5.0%	6.1%	5.1%	5.4%	5.6%
avg shortfall	2.0%	2.2%	2.1%	1.6%	1.8%	1.7%	1.5%	1.6%	1.6%	1.5%	1.4%	1.4%
stdev of 36-month returns	4.9%	4.5%	4.2%	3.9%	4.5%	3.6%	3.6%	3.5%	4.6%	3.5%	3.8%	3.9%
monthly stdev	8.9%	4.0%	4.3%	5.5%	9.0%	4.7%	5.4%	5.7%	9.0%	5.7%	6.4%	6.6%
max drawdown	20.8%	5.1%	5.4%	9.4%	20.8%	9.4%	12.5%	13.9%	21.7%	13.2%	16.8%	17.5%
Shapre ratio (36-month stdev)	61.7%	36.2%	37.9%	61.5%	65.4%	55.9%	62.5%	63.2%	73.4%	65.9%	70.2%	72.0%
Sharpe ratio (monthly stdev)	34.3%	40.8%	37.1%	43.6%	32.9%	42.5%	42.0%	39.2%	37.2%	40.5%	41.5%	42.9%
	Maintain AA for 36-months											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
avg return. p.a.	13.1%	11.5%	11.9%	12.3%	13.2%	12.0%	12.3%	12.3%	13.3%	12.4%	12.5%	12.6%
avg real return p.a.	5.4%	3.7%	4.2%	4.6%	5.5%	4.3%	4.5%	4.6%	5.6%	4.7%	4.8%	4.9%
avg shortfall	2.1%	2.9%	2.4%	2.2%	1.9%	2.2%	2.1%	2.1%	1.8%	2.0%	2.0%	1.9%
stdev of 36-month returns	4.2%	3.8%	3.4%	3.3%	4.2%	3.0%	3.1%	3.1%	4.1%	3.1%	3.3%	3.4%
avg monthly stdev	8.7%	3.7%	4.1%	5.3%	8.8%	4.7%	5.3%	5.6%	8.9%	5.6%	6.3%	6.5%
avg max drawdown	12.0%	6.1%	6.0%	7.7%	12.0%	6.6%	7.5%	7.9%	12.3%	7.9%	9.1%	9.3%
Shapre ratio (36-month stdev)	51.3%	14.2%	30.4%	42.4%	54.5%	37.3%	44.0%	44.0%	58.1%	47.1%	48.5%	49.7%
Sharpe ratio (monthly stdev)	25.1%	14.8%	25.5%	26.2%	25.8%	24.3%	25.5%	24.7%	26.7%	26.1%	25.2%	26.1%
	Average asset allocation											
	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3	no regime	regime 1	regime 2	regime 3
SA equities	39%	3%	10%	15%	43%	22%	24%	26%	44%	28%	30%	31%
SA bonds	9%	27%	28%	40%	12%	14%	19%	20%	16%	14%	19%	20%
SA cash	12%	43%	37%	17%	15%	48%	37%	35%	13%	43%	33%	32%
Global equities	4%	12%	8%	4%	1%	1%	0%	0%	1%	0%	0%	0%
Global bonds	35%	14%	18%	24%	30%	16%	19%	19%	27%	15%	18%	17%
Total equities	44%	15%	18%	19%	44%	22%	24%	26%	45%	28%	30%	31%
Total offshore	39%	26%	26%	28%	31%	16%	19%	19%	28%	15%	18%	17%

Table 50: Standard deviation over horizon optimal porfolios for a required real return of 6% and investment horizon of 36 months, for three different block lengths (one, 12, 24 months), backtested with returns data from 1972 to 2017.

8. REFERENCES

Ackora-Prah, J., Gyamerah, S.A. & Andam, P.S. 2014. A heuristic crossover for portfolio selection. *Applied Mathematical Sciences*. 8(65):3215-3227.

Aït-Sahalia, J. & Brandt, M.W. 2001. Variable Selection for Portfolio Choice. *The Journal of Finance.* 56(4):1297-1351.

Ang, A. & Chen, J. 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics.* 63(3):443-494.

Ang, A. & Bekaert, G. 2002. International Asset Allocation with Regime Shifts. *The Review of Financial Studies.* 15(4):1137-1187.

Ang, A. & Bekaert, G. 2004. How do Regimes Affect Asset Allocation? *Financial Analysts Journal*. 60(2):86-89.

Ang, A. & Timmermann, A. 2012. Regime Changes and Financial Markets. *Annual Review of Financial Economics.* 4(1):313-337.

Bae, G. Il., Kim, W.C. & Mulvey, J.M. 2014. Dynamic asset allocation for varied financial markets under regime switching framework. *European Journal of Operational Research.* 234(2):450-458.

Bauer, R., Haerden, R. & Molenaar, R. 2004. Asset Allocation in Stable and Unstable Times. *The Journal of Investing.* 13(3):72-80.

Black, F. & Scholes, M. 1973. The pricing of options and corporate liabilities. *The Journal of Political Economy.* 81(3):637-654.

Bossaerts, P. & Hillion, P. 1999. Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies.* 12(2):405-428.

Bradfield, D., Dugmore, B. & Gopi, Y. 2006. A new measure of portfolio diversification. Research Paper, Cadiz.

Brereton, R.G. 2015. The Mahalanobis distance and its relationship to principal component scores. *Journal of Chemometrics.* 29(3):143-145.

Carhart, M.M. 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance*. 52(1):57.

Carlstein, E. 1986. The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence. *The Annals of Statistics.* 14(3):1171-1179.

Carlstein, E., Do, K., Hall, P., Hesterberg, T. & Kunsch, H.R. 1998. Matched-block bootstrap for dependent data. *Bernoulli*. 4(3):305-328.

Castro, R. No date, The Empirical Distribution Function and the Histogram, lecture notes, Eindhoven University of Technology. Available: https://www.win.tue.nl/~rmcastro/2WS17/files/ecdf_hist.pdf

Çelikyurt, U & Özekici, S. 2007. Multiperiod portfolio optimization models in stochastic markets using the mean-variance approach. *European Journal of Operational Research.* 179(1):186-202.

Chang, T.J., Yang, S.C. & Chang, K.J. 2009. Portfolio optimization problems in different risk measures using genetic algorithm. *Expert Systems and Applications.*

Chen, S.H. (ed.) 2004. *Genetic Algorithms and Genetic Programming in*

Computational Finance. Berlin: Springer.

Cheung, B.K., Tang, G.Y., Kong, H. & De Montreal, E.P. 1996. Genetic Algorithms in Multi-Stage Portfolio Optimization System. *Optimization.* 1995. 36(7):10529-10537.

Chow, G., Jacquier, E., Kritzman, M. & Lowry, K. 1999. Optimal Portfolios in Good Times and Bad. *Financial Analysts Journal.* 55(3):65-73.

Cont, R. 2001. Empirical properties of assets returns: Stylized facts and statistical issues. *Quantative Finance.* 1(2):1-14.

Cosslett, S.R. & Lee, L.F. 1985. Serial correlation in latent discrete variable models. *Journal of Econometrics.* 27(1):79-97.

Dimson, E., Staunton, M. & Marsh, P. 2001. *Triumph of the Optimist: 101 years of Global Investment Returns*. Princeton: Princeton University Press.

Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics.*7(1):1-26.

Fama, E.F. & French, K.R. 1988. Permanent and Temporary Components of Stock Prices. *Journal of Political Economy.* 96(2):246-273.

Fama, E.F. & French, K.R. 1993. Fama-French APT three-factor model. *Journal of Financial Economics*. 33(1): 3-56.

Fama, E.F. & French, K.R. 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics.* 105(3):457-472.

Firer, C. & McLeod, H. 1999. Equities, Bonds, Cash and Inflation: Historical Performance in South Africa 1925 to 1998. *Investment Analysts Journal.* 50 (April):7-28.

Fraser, A. M. 2008. *Hidden Markov Models and Dynamical Systems*. Philadelphia: Society for Industrial and Applied Mathematics Press.

Fu, J., Wei, J. & Yang, H. 2014. Portfolio optimization in a regime-switching market with derivatives. *European Journal of Operational Research.* 233(1):184-192.

Gray, S. 1996. Modeling the conditional distribution of interests in a regime-switching process. *Journal of Financial Economics.* 42(1):27-62.

Guidolin, M. & Timmermann, A. 2005. *Optimal Portfolio Choice under Regime Switching , Skew and Kurtosis Preferences*. Federal Reserve Bank of St. Louis Working Paper Series.

Guidolin, M. & Timmermann, A. 2007. Asset Allocation under Multivariate Regime Switching. *Journal of Economic Dynamics and Control.* 31(11):3503-3544.

Guidolin, M. & Timmermann, A. 2008. International asset allocation under regime switching, skew, and kurtosis preferences. *The Review of Financial Studies.* 21(2):889-935.

Hall, P. 1985. Resampling a coverage pattern. *Stochastic Processes and their Applications.* 20(2):231-246.

Hamilton, J. 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica.* 57(2):357-384.

Hamilton, J.D. 1996. Specification testing in Markov-switching time-series models. *Journal of Econometrics.* 70(1):127-157.

Honda, T. 2003. Optimal Portfolio Choice for Unobservable and Regime Switching Mean Returns. *Journal of Economic Dynamics and Control.* 28(1):45-78.

Ibbotson, R. G. & Kaplan, P.D. 2000. Does Asset Allocation Policy Explain 40, 90, or 100 Percent of Performance? *Financial Analysts Journal.* 56(10)*:*26-33.

Kaya, H., Lee, W.A.I. & Pornrojnangkool, B. 2010. Regimes: Nonparametric Identification and Forecasting. *Journal of Portfolio Management.* 36(2):94-105.

Koijen, R.S.J., Rodriguez, J.C. & Sbuelz, A. 2009. Momentum and mean reversion in strategic asset allocation. *Management Science.* 55(7):1199-1213.

Kondlo, M. 2016. *A framework for regime identification and asset allocation*. MA Thesis in Statistical Sciences. University of Cape Town. Available: https://open.uct.ac.za/bitstream/handle/11427/20475/thesis_sci_2016_kondlo_mpumelelo.pdf?sequence=1 [2018, 6 November].

Kritzman, M., Page, S. & Turkington, D. 2012. Regime shifts: Implications for dynamic strategies. *Financial Analysts Journal.* 68(3):22–39.

Kunsch, H.R. 1989. The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics* 17(3):1217-1241.

Lahiri, S.N. 2003. *Resampling Methods for Dependent Data*. Berlin: Springer Science & Business Media.

Lam, K., So, M. & Li, W.K. 1998. A Stochastic Volatility Model With Markov Switching. *Journal of Business and Economic Statistics.* 16(2):244-253.

Leinweber, D. & Arnott, R.D. 1995. Quantitative and Computational Innovation in Investment Management. *The Journal of Portfolio Management.* 21(2):8-15.

Longin, F. & Solnik, B. 2001. Extreme correlation of international equity markets. *The Journal of Finance*. 56(2):649-676.

Mahalanobis, P.C. 1927. Analysis of Race-Mixture in Bengal. *Journal of the Asiatic Society of Bengal.* 23:301-333.

Markowitz, Harry M. 1952. Portfolio Selection. *The Journal of Finance.* 7(10):77-91.

Merton, R.C. 1969. Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics.* 51(3):247.

Merton, R.C. 1973. Theory of rational option pricing. *The Bell Journal of Economics and Management Science.* 4(1):141-183.

Meucci, A. 2008. *Risk and Asset Allocation.* Berlin: Springer Finance.

Mulvey, J.M., Rosenbaum, D.P. & Shetty, B. 1997. Strategic financial risk management and operations research. *European Journal of Operational Research.* 97(1):1-16.

Munro, B. & Silberman, K. 2008. Optimal Asset Allocation in Different Economic Environments. *Cadiz Securities Quantitative Research Report.*

Nystrup, P. 2014. *Regime-Based Asset Allocation: Do Profitable Strategies Exist ?* MA Thesis. Technical University of Denmark. Available: http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6808/pdf/imm6808.pdf [2018, 6 November].

Nystrup, P., Hansen, B.W., Madsen, H. & Lindström, E. 2015. Regime-Based Versus Static Asset Allocation: Letting the Data Speak. *Journal of Portfolio Management.* 42(1):103-109.

Nystrup, P., Madsen, H. & Lindström, E. 2017. Long Memory of Financial Time Series and Hidden Markov Models with Time-Varying Parameters. *Journal of Forecasting.* 36(8):989-1002.

Politis, D.N. & Romano, J.P. 1992. A circular block-resampling procedure for stationary data. In *Exploring the Limits of Bootstrap*. LePage, R. & Billard, L., eds. New York: John Wiley. 263-270.

Politis D.N. & Romano, J.P. 1994. The stationary bootstrap, *Journal of the American Statistical Association.* 89(428):1303-1313.

Poterba, J.M. & Summers, L.H. 1988. Mean reversion in stock prices: Evidence and Implications. *Journal of Financial Economics.*22(1):27-59.

Psaradakis, Z. & Spagnolo, N. 2003. On the determination of the number of regimes in Markov-switching autoregressive models. *Journal of Time Series Analysis.* 24(2):237-252.

Sa-Aadu, J., Shilling, J.D. & Tiwari, A. 2005. *Portfolio Performance and Strategic Asset Allocation Across Different Economic Condition*s. Working paper, University of Iowa.

Schwert, G.W. 1989. Why does stock market volatility change over time? *Journal of Finance*. 44(5):1115-1153.

Simonde de Sismondi, J.C.L. 1819. *New Principles of Political Economy, vol 1*. 1990 Edition. Translated and annotated by Richard Heyse, and with a foreword by Robert Heilbroner. New Brunswick: Transaction Publishers.

Spierdijk, L., Bikker, J.A. & Van den Hoek, P. 2012. Mean reversion in international stock markets: An empirical analysis of the 20th century. *Journal of International Money and Finance.* 31(2):228-249.

Timmermann, A. 2000. Moments of Markov switching models. *Journal of Econometrics.*96(1):75-111.

Tsay, R.S. 2005. *Analysis of financial time series*, *vol. 543*. New York: John Wiley & Sons.

Tu, J. 2010. Is Regime Switching in Stock Returns Important in Portfolio Decisions? *Management Science.*56(7):1198-1215.

Turing, A.M. 1950. Computing machinery and intelligence. *Mind.* 59(October):433.

Turner, C. & Han, J. 2009. Portfolio Optimization under Time-Varying Economic Regimes. Paper presented at Stanford University. Available: http://cs229.stanford.edu/proj2009/HanTurner.pdf [2018, 6 November].

Van der Vaart, A.W. 2000. *Asymptotic Statistics*. (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge: Cambridge University Press.

Van Wyk de Vries, E., Gupta, R. & Van Eyden, R. 2014. Intertemporal portfolio allocation and hedging demand: An application to South Africa. *Journal of Business Economics and Management.* 15(4):744-775.

9. PLAGIARISM DECLARATION

<table>
<tr><td>Name</td><td colspan="3">Barry van Zyl</td></tr>
<tr><td>Student No:</td><td colspan="3">VZYBAR001</td></tr>
<tr><td>Tel numbers:</td><td>021 524 4430</td><td>Email address:</td><td>barryvz@gmail.com</td></tr>
<tr><td>Word count:</td><td>47 650</td><td>No. of pages</td><td>165</td></tr>
<tr><td>Dissertation Title:</td><td colspan="3">The Optimal Asset Allocation for South African Real Return Investors</td></tr>
<tr><td>Name of Supervisor/s:</td><td colspan="3">Prof David Bradfield</td></tr>
<tr><td colspan="4">DECLARATION:
1. I am presenting this dissertation in ~~FULL~~/PARTIAL fulfilment of the requirements for my degree.
2. I know the meaning of plagiarism and declare that all of the work in the dissertation, save for that which is properly acknowledged, is my own.
3. I hereby grant the University of Cape Town free licence to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever of the above dissertation.</td></tr>
<tr><td>Signature</td><td>Signature Removed</td><td>Date:</td><td>2018-12-18</td></tr>
</table>