A FRAMEWORK FOR REGIME IDENTIFICATION AND ASSET ALLOCATION

BY: MPUMELELO KONDLO
SUPERVISOR: PROF. DAVID BRADFIELD

Masters Thesis in Statistical Science
Department of Statistical Sciences
Univers ity of Cape Town
March 2016

ABSTRACT

The purpose of this thesis is to examine a regime-based asset allocation strategy and evaluate whether accounting for regime-dependent risk and return of asset classes provides any significant improvement on portfolio performance. The South African market and economy are considered as a proxy for the analysis. Motivation of this thesis stems from the growing body of research by practitioners devoted to models that are reflective of the interdependency between financial assets and the real economy. The asset classes under consideration for the analysis are domestic and foreign cash, domestic and foreign bonds, domestic and foreign equity, inflation linked bonds, property, gold and commodities.

In order to evaluate the performance of the regime-based strategy, this thesis proposes a framework based on Principal Component Analysis and Fuzzy Cluster Analysis for regime identification and asset allocation. The performance of the strategy is tested against two strategies that are not cognizant of regime changes. These are an equally weighted portfolio and a buy-and-hold strategy. Furthermore, relative performance analysis was performed by comparing the regime-based strategy proposed in this thesis against the Alexander Forbes Large Manager Watch Index. Due to data limitations, the analysis is done on an in-sample basis without an out-of-sample testing.

The results from the analysis showed the extent of outperformance of the proposed regime-based strategy relative to an equally weighted strategy and a buy-and-hold strategy. These results were consistent with existing literature on regime-based strategies. Furthermore, the results provided strong motivation for the use of the regime identification framework together with tactical asset allocation proposed in this thesis.

ACKNOWLEDGMENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACRONYMS

x

Part I

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

"By the Law of Periodical Repetition, everything which has happened once must happen again and again and again - and not capriciously, but at a regular period, and each thing in its own period, not another's and each obeying its own law . . . the same Nature which delights in periodical repetition in the skies is the Nature which orders the affairs of the earth. Let us not underrate the value of that hint"

- Mark Twain

The understanding of cyclical trends in any system, if they exist, provides an attractive and informative framework in which rational decisions can be based. Ranging from biogeochemical cycles in earth's natural sciences to thermodynamic cycles in engineering and the cardiac cycle of the human body, the study of cycles has provided an exceptional framework for decomposing and understanding complex systems. The financial market is no exception to cyclical phenomena. Indeed, the procyclical nature of financial markets, though not well defined (Borio, 2012), presents a vital stepping stone for long-term portfolio managers upon which to rationalise their asset allocation strategies.

The so called "boom – bust" periods that have repeatedly occurred in financial markets has led to a growing body of literature dedicated to understanding and modelling these occurrences. However, due to the irregularities inherent in financial cycles and the constantly changing determinants that precipitates each cycle, researchers have had little success in establishing a formalised "cause – effect" framework that is robust over time (Borio, 2012). Such complexity has led to many cross-discipline researchers seeking ways to indirectly study financial cycles in order to gain a competitive edge in their asset allocation frameworks. Most notably, the increasing interdependency between financial markets and the real economy has given rise to the reconciliation of financial and business cycles to help position investment portfolio's (Forest, Orpiszewski, Péters and Kojo, 2014).

Forest et al (2014) demonstrate in their work how monitoring cyclical changes in the business cycle can provide a significant improvement in the asset allocation decision. However, other studies have shown that the relationship between financial and business cycles easily breaks down (Avouyi-Dovi and Matheron, 2004; Classens, Kose and Terrones, 2011). While business cycles result from a mere fluctuation in business activities, financial cycles are characterised by self-reinforcing interaction between perception of value, risk and other

investor subjectivities which are translated into boom-bust periods (Borio, 2012). These fundamental differences between financial and business cycles result in large drawdowns on asset allocation strategies based on these relationships.

Nevertheless, the established importance of the macro environment inherent in business cycles provides an important lense through which we can analyse financial markets for asset allocation purposes. There are two key points of interest that can be deduced from studies relating to financial and business cycles. Firstly, even though the nature of financial cycles differ from time to time, the general path which they follow remains the same (Classens et al, 2011; Borio, 2012). Secondly, and perhaps more importantly, there exists a set of economic conditions that characterises different investment environments which subsequently influences the behaviour of the different assets within the financial cycle (Eychenne and Martinetti, 2011). These cyclical economic conditions or, rather, economic regimes highlight periods of consistent uniqueness in the joint and relative behaviour of investment assets.

This motivates for a regime-based framework for decision-making which can account for volatility and correlation patterns of investment assets over the financial cycle. Furthermore, the presence of regime changes has posed serious challenges to traditional asset allocation processes due to their inability to account for time-varying risk propelled by these shifts.

This thesis examines a regime-based asset allocation strategy which takes account of regime-dependent risk and returns to establish whether any improvement in portfolio performances over static allocation is warranted. The remainder of this thesis is organized in the following manner: Chapter 2 of this section provides the literature review, thesis statement and delimitations of the study. This is followed by Part 2 which is composed of Chapter 3 and 4. Chapter 3 provides data analysis for the data used in this study while Chapter 4 entails the methodology that will be followed for the regime identification and portfolio optimization process for asset allocation. The thesis ends with Part 3 composed of chapter 5 and 6 which provides the results, conclusion and recommendation of further work in this area.

2

LITERATURE REVIEW AND THESIS STATEMENT

2.1 LITERATURE REVIEW

The art of asset allocation has historically been considered as one of the most important decisions in an investment process (Bulla, Mergner, Sesboue, and Chesneeau, 2010). It is argued in academic literature that the asset allocation decision attributes 90% of the ultimate portfolio performance. (Etula and Farshid, 2012). Traditionally, the asset allocation process has been approached as a bottom-up exercises where investors would seek to allocate their funds based on the risk-return profile of individual assets. This approach can be dated back to the early works of Markowitz (1952).

Markowitz (1952) established the Mean-Variance framework based on the analysis of expected returns and risk of assets within the investment universe. Ceteris paribus, the asset allocation process under the Markowitz framework involved choosing asset classes whose expected returns maximized the risk-adjusted returns of the overall portfolio. The novel approach suggested by Markowitz (1952) was based on a number of assumptions which presented some short-comings when relaxed for practical purposes. Amongst these, two limitations are highlighted from a long-term portfolio management point of view.

Firstly, the Markowitz (1952) framework was based on a single period model of investment (Kaplan, 1998). This issue results in static portfolios that are prone to time-variation. For long-term portfolio management, this would imply more frequent rebalancing which may lead to higher costs. Secondly, and perhaps more importantly for this study, the asset allocation process appeared as a "black box" process. That is, the process does not provide the means for practitioners to actively engage with the allocation process through expressing their subjective views. Thus, small changes in the investment environment that causes deviations from forecasted point estimates can result in extreme portfolios.

Black and Litterman (1991) were amongst the first authors to provide a framework that allowed for the specification of investor views which could be blended with historical information of asset classes (Walter, 2007). The Black – Litterman model improved on the Markowitz framework by merging the equilibrium Capital Asset Pricing Model with subjective views on asset returns (Walter, 2007). The inclusion of subjective views within the return estimation process provided a platform where investors could account for changes in their investment environment. Post 1992, a taxonomy of equivalent models were

5

presented in academic literature. This includes models developed by Bevan and Winkelmann (1998), He and Litterman (1999), Satchell and Scowcroft (2000), Herold (2003), etc.

Krishnan and Mains (2005) extended the Black and Litterman model to a Two – Factor Black – Litterman model by adding an extra factor which was uncorrelated to the market. The authors demonstrated that the inclusion of the extra "recessionary" factor to the standard Black-Litterman model impacted the expected returns computed from the model. This result signified the importance of the macro-environment on the resultant returns of asset classes, hence the need to consider macro-economic scenarios.

Almgren and Chriss (2004) modified the Markowitz portfolio selection method by proposing an optimisation model that was based on directional signals rather than the point estimates of asset returns. Similar to Black and Literman (1991), the model proposed by Almgren and Chriss (2004) allowed investors to incorporate their subjective views in their asset allocation process. However, unlike the framework proposed by Markowitz (1952) and Black and Litterman (1992), the use of directional signals as input into the optimisation implied that the model was not subject to changes in point estimates of asset returns.

The aforementioned models improved on some of the short comings presented by the original mean-variance framework but remained static in nature. Authors such as Steinbach (1999) and Araujo and Costa (2008) provided extensions of the single period Markowitz framework to a multi-period framework while Ledoit and Wolf (2003) proposed covariance shrinkage methods which reduces estimation error in the models. However, the formulation of the models remained subject to time-varying risk that arises as a result of regime shifts. Empirical evidence has shown that the returns of asset classes are subject to regime shifts within the investment environment. This includes the works of Bulla, Mergner, Sesboue and Chesneeau (2010), Amenc, Malaise, Martellini and Sfeir (2003) and Ang and Bekaert (2002). This realisation gave birth to the development of regime-based asset allocation strategies.

Bulla et al (2010) conducted a study that examined the profitability of a regime based asset allocation strategy relative to buy-and-hold strategies with transaction costs. Although the authors did not focus on individual asset performances in different investment environments, they did show that regime-based strategies outperform the buy-and-hold approaches. In support of this result, Ang and Bekaert (2002) illustrated in their work how the presence of regimes can be incorporated in an asset allocation strategy. The authors concluded that expected returns and volatilities vary significantly with time. Furthermore, they found that in a volatile environment, correlation amongst equity returns increase significantly and asset prices fell.

Amenc et al (2003) provided an analysis of market-neutral portfolios that were aimed at producing absolute returns over the full business cycle. The authors found that style indexes performed differently in different economic environments due to the different economic and financial risk factors they were exposed to. Although no particular regimes were defined, the authors postulated that one of the approaches that is viable for return forecasting is to first forecast the value of the economic variables.

Gildfish and Quandt (1973) were amongst the first to introduce time–varying parameter models in the field of finance through Markov Switching regression. The authors focused on disequilibrium that persisted in the housing market. However, in the area of macro finance, such models gained traction after the seminal work of Hamilton (1989). In his seminar, Hamilton (1989) modelled time series data using a Two–State Markov Switching model. This provided the foundation of regime-based allocation where the transitioning from one regime to the next was governed by a latent Markov Process.

Sa-Aadu, Shilling and Tiwari (2005) used a Markov Regime – Switch framework to examine portfolio allocation under time-varying investment environments where returns and volatilities of asset classes vary with economic regimes. Under this study, the authors classified the different economic regimes as changes in the per capita consumption growth rate. They found that, over time, the ability of different asset classes to improve portfolio performance and hedge against adverse shocks in consumption growth was time dependent, and that the hedging role from the different asset classes alternated between them. Furthermore, the authors found that while a group of assets provided outperformance during times of positive consumption growth, a different set improved portfolio performance during mean consumption periods, while other asset classes provided strong hedging against deterioration in consumption opportunities.

Bae, Kim and Mulvey (2003) addressed the time-varying risk associated with financial assets by extending the Markowitz framework to account for the changing nature of the covariance matrix under different market environments. In order to account for the different conditions in the market, hence regimes, a Markov Regime Switching frame-work that uses Hidden Markov (HMM) processes was implemented. The different market regimes in this model were defined by the different states assumed by the Markov process. The model used in this paper was a four state HMM. Under the different states, the authors found that the three asset classes under consideration performed differently. More importantly, the authors established that the correlation structure in each state was different and no single asset was superior in performance in all four states.

Using Markov Chain Monte Carlo (MCMC) methods, Ammann and Verhofen (2005) constructed a multivariate regime switch model

to analyse time-varying risk premia and their implications for portfolio choice. In their study, the authors found two distinct regimes in their data. These regimes had different means, volatilities and correlations. The first regime was characterized by high volatilities and correlations with low returns while the other regime was characterised by low volatilities with high returns and occurred about 75% of the time. They also found the transition between extreme regimes was less likely than transitions between more similar regimes. Similar work by Fama and French (1998) showed how value premiums and size premiums could be related to the growth in the economy.

Yin and Yu Zhou (2003) interrogated a class of discreet-time mean-variance portfolio selection models and illustrated the relationship of these models with continuous-time models. In this paper, the authors used a Markov-Modulated geometric Brownian motion framework to formulate a portfolio selection model. The striking feature in their approach was the use of latent state distribution modes for the different regimes the market may go through. In their explanation, transitioning between the different modes, hence regimes, reflected the different states of the economy, investor mood and economic factors. This further highlighted the importance of the macro - environment when considering the asset allocation process.

Guidolin and Timmermann (2005) proposed a multi-period portfolio allocation approach that can be used in the presence of Regime Switching models. Their model was derived under the assumptions that the distribution of the asset returns can be modelled using a regime switch process. Similarly, Araujo and Costa (2008) considered a Multi-Period Markowitz model with Markov switching parameters. In this model, the latent regimes were drawn from a multinomial distribution that assumes no autocorrelation. In their study, these authors considered random shocks to the asset classes collectively and to each asset class on an individual level. A Bayesian portfolio selection method with Markov Switches was then implemented for asset allocation. The resultant portfolio composition varied significantly over the time period under consideration and was very responsive to the shocks induced on the asset classes.

The Empirical results presented by the class of models under the umbrella of Markov Regime Switch framework provides a strong case for regime-based asset allocation. However, there are concerns that arises from this family of models. Firstly, Markov Regime Switch models are parametric in nature (Kneib, Langrock and Sohn, 2015). In determining the parameters of the model, the user is required to assume an observation generating probability density function. The authors that have been reviewed thus far made the assumption of the Gaussian probability distribution in their models. This assumption has been shown not to hold in numerous papers (Mittnik, 1963; Fama, 1965; Biglova et al., 2004). Furthermore, the use of a Gaussian

distribution necessarily assumes a zero correlation structure in the data (Kneib et al, 2015). This also does not always hold in practice as shown in numerous studies that correlations tend to increase during bad times.

Secondly, the criteria used in these models to define the different regimes are not necessarily intuitive and informative (Sa-Aadu et al, 2005). The most common method used to define the different regimes in the Bayesian Information Criteria (BIC). In this case, the resultant regimes defined do not provide any information on the structure of the different regimes (Sa-Aadu et al, 2005). Furthermore, since these regimes are derived endogenously from the asset class data being examined, it ignores the direct impact of exogenous factors that may arise from the underlying macroeconomic environment. There was no direct investigation of the impact of the investment environment as defined by economic variable on the behaviour of asset classes. The second leg of regime based models provide a closer interrogation of the macro-environment on the behaviour of asset classes.

Eychenne and Martinetti (2011) proposed a systematic approach in estimating long run asset returns. Their method was based on the notion of fair value in asset prices where the long run asset returns are conditional on the long-run path of the real economy. In justifying their model, the authors initiated their approach by highlighting the short-comings of the unconditional regime switch models based on historical figures as a guide for the future. The major drawdown mentioned was that these models assume that the long-run path of the economy is stationary and constant. This caused underperformance in portfolios when the economy goes through cycles and assets respond to such cycles. To account for these regime changes in the macro-environment, the authors considered a return estimation method that is conditional on global economic scenarios. The long-run economic scenarios were obtained using two fundamental economic indicators, namely output as measured by real GDP and Inflation. Even though the study conducted by these authors was not aimed at analysing the time-varying risk of assets presented by regime changes in the economy, they did find that conditioning asset return expectations on macroeconomic variables provided a more plausible strategic asset allocation.

Kollar (2003) focused on the building blocks of an asset allocation strategy based on the interaction between the science of macroeconomics and that of finance. Similar to Eychenne and Martinetti (2011), the author used inflation and GDP growth to define the economic regimes. Before introducing the framework, the authors start by pointing out some stylized facts that hinder traditional asset allocation approaches. Reiterating these stylised factors, the authors showed through an observational study that risk premiums were not constant over time. Furthermore, they showed that over time, asset class re-

Table 1: Different Economic Environments

Economic Environments		Inflation	
		Rising	**Falling**
Real GDP Grouth	**Rising**	*Expansion*	*Correction*
	Falling	*Stagflation*	*Contraction*

turns exhibited variation in their correlation structure. Drawing from these stylised factors, the author found that asset class behaviour differed significantly during different phases of the economic cycle. Furthermore, Kollar (2003) established that no single asset class exhibited outperformance across all defined economic conditions.

Chalmer, Kaul and Phillip (2011) conducted a study that analyses the behaviour of US mutual funds with respect to their asset allocation decision. One of the questions addressed by the study was whether or not mutual fund managers reacted to changing economic conditions when making their asset allocation decisions. To account for the different economic conditions, the authors used a PCA proxy for economic variables and a set of financial market proxies. Similar to Sa-Aadj et al (2005), the authors used a Markov Switch framework and regression to analyse portfolio rebalancing under different economic conditions. They found that there was a consistent portfolio risk adjustment in response to shifting economic conditions. Most notably, there was a consistent "flight – to – safety", away from risky equity funds to low risk money market funds, during periods of economic downturn. On the other hand, when expectations of a favourable economy rise, investors direct their funds away from money market funds towards risky equity funds.

The common methodology that is followed by Chalmer, Kaul and Phillip (2011) and Kollar (2003), in analysing the performance of the different regimes was to construct the different regimes separately. The framework for regimes defined in the work of Eychenne and Martinetti (2011) and Kollar (2003) can be summarised in Table 1 above.

The idea behind this method was to find periods that satisfy these conditions in the period under consideration. The data is then spliced and the behaviour of the different asset classes within these periods is then analysed. The results from Eychenne and Martinetti (2011) and Kollar (2003) certified that the asset classes indeed exhibited different behaviour within these periods.

The use of the GDP/Inflation framework for regime construction makes the assumption that the regimes assumed by the economy can be derived from the behaviour of these variables. This is not necessarily always the case. Academic research has shown that there are other variables that are significant in explaining economic activities. This includes the works of Mueller (2007) and Gilchrist and Zakrajsek (2011)

who demonstrated a consistent relationship between credit spreads and fluctuations in the real economy and the business cycle. Bleaney, Mizen and Veleanu (2012) further showed that the credit spreads can be used to predict economic activities.

Pretorius and Venter (2004) provided 29 composite leading and co-incidental business cycle indicators for the South African economy. More than half of these variables were macroeconomic variables. Similarly, a study by Walt (1982) showed how these economic indicators are identified, classified and adjusted through time. He further provided some statistical results that measures the resultant performance of the economy as captured by these variables. This demonstrates that regime shifts in the economy arises as a result of a combination of variables, hence considering only two variable may be restrictive.

In an attempt to account for more variables in the regime classification, Chalmer et al (2011), mentioned earlier, used a regression approach. In this method, Chalmer et al (2011) makes the assumption of a linear relationship between the returns of the asset classes and the macroeconomic variables. This assumption is a very strict one as there is no proof that is provided that certifies this relationship. Furthermore, each time a new data point is added to the data set, the coefficients of the regression model would need to be re-estimated.

Blitz and Vliet (2009) used four different variables in their study of dynamic strategic asset allocation. The aim of their study was to provide a regime-based asset allocation method that enhances portfolio returns while stabilising portfolio risk across different economic regimes. The authors considered a sixty year period spanning from 1948 to 2007. Their study covered the major three asset classes namely equity, bonds and cash together with sectors grouped into small cap, value, growth, credits and commodities. In defining the different economic regimes, the authors used two macro factors; the seasonally adjusted U.S ISM manufacturing survey production index and the seasonally adjusted U.S unemployment rate; and two market factor – credit spreads and earnings yields. Each regime was then defined by a combination of the level of these factors. The regime-based asset allocation strategy was then compared to the performance of static and tactical asset allocation strategies. The findings from these authors was that the more dynamic approach of considering regime changes outperformed the other strategies.

In their study, Hun and Turner (2010) used Principal Component (PC) decomposition in order to account for the number of variables that are perceived to affect the economic environment. The aim of their study was to assess whether or not accounting for macroeconomic conditions can lead to better investment allocation amongst different sectors. The authors considered two approaches for regime identification. The first approach was based on cluster analyses where the k-means method was applied to the reduced data from the PCA
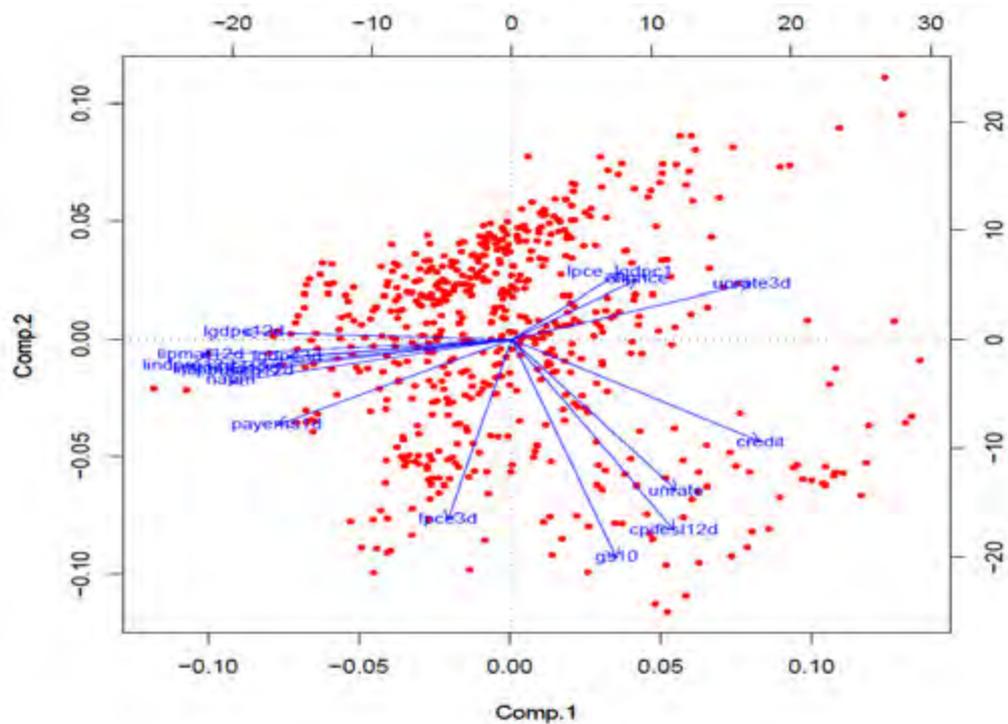
Figure 1: Biplot relationship of 1st and 2nd PC
Source: Turner and Han (2004)

decomposition. The authors rationalised the use of PCA by illustrating how the PC's captured the essential information embedded in the data set. Before the PCA was applied, the authors provided a visual display of the structural behaviour of the economic variables. These results are presented in Figure 1 above with the names of the variables used in the chart given in the appendix.

The main reason highlighted by the authors for such a pre-analysis of the data was to ensure that the clustering of the chosen variables occurs in a sensible way. In so doing, results that would be derived from clustering techniques such as the k-means would be consistent in their interpretation. To motivate for the use of the reduced data via the PCA, the authors showed that major events that transpired within the economy are captured in the scores derived from the PCA decomposition. Figure 2 on the next page provides a snap-shot of results from the paper. This shows events captured by the first PC in the data set used by the authors.

In this case, the authors demonstrated that it is plausible to use a set of these PC's for the analysis rather than using the entire data set. These PC's were able to capture major economic trends and various aspects of the economy such as inflationary concerns and credit – tightening periods (Hun and Turner, 2010).

In this thesis, we draw from the works of Blitz and Vliet (2009) and Hun and Turner (2010). There are two additions that this thesis
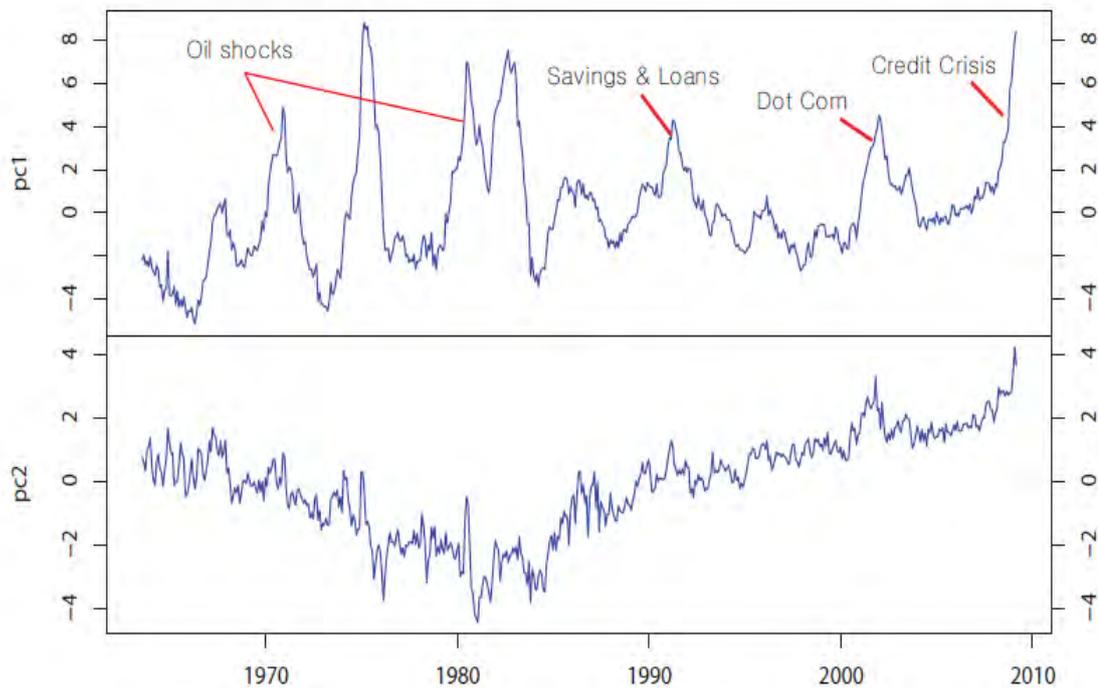
Figure 2: Projection   of   Economic   State   onto   First   Two   PC's
Source: Turner and Han (2004)

makes which has not been analysed by the aforementioned papers. Firstly, similar to Hun and Turner (2010), we use a PCA decomposition on a set of macroeconomic variables to reduce the dimensionality of the data. However, contrary to Hun and Turner (2010), we use a Fuzzy clustering method as opposed to Markov Regime-Switch models to establish our regimes. After establishing the regimes, subsequent analysis of asset classes is made following the work of Blitz and Vliet (2009) who similarly defined the regimes separately, but limiting their economic variables to four, our study instead considers a range of variables.

2.2 THESIS STATEMENT

The purpose of this thesis is to examine a regime-based asset allocation strategy for long-term portfolio management and to assess whether accounting for regime-dependent risk and return of assets provides any significant improvement on portfolio performance. It will be considered whether the outperformance of such a strategy demonstrated in previous studies on regime-based asset allocation is worthwhile within an emerging economy using South Africa as a proxy.

The relevance of such a study is supported by the increased number of articles devoted to the subject by practitioners in search of more realistic asset allocation strategies that reflects the random market environment. This includes unpublished research by Munro and Silberman (2008) from Cadiz Asset Management, Sheikh and Sun (2011) from J.P. Morgan Asset Management, Davis, Aliaga-Díaz and Patterson, (2011), Briere and Signori (2012) from Amundi Asset Management.

Contrary to the body of literature on regime-based asset allocation, this thesis consider a broader base of asset classes. The asset classes considered are local and foreign cash, local and foreign bonds, local and foreign equity, property, gold, commodities and inflation linked bonds. It is worth mentioning that while the other studies considered a subset of these asset classes, no study was found that considered inflation linked bonds as a separate asset classes. Furthermore, one novelty that is added by this thesis is the proposed approach to regime classification for asset allocation. To the best of our knowledge at the time of writing, no literature was found that combined Principal Component Analysis with Fuzzy Cluster Analysis for regime classification purposes for asset allocation. It is also worth mentioning that while other studies considered a subset of these asset classes, no study was found that considered inflation linked bonds as a separate asset class.

The data contains asset monthly returns spanning the period from 1993 to 2014. Due to the limited data points, the full data set was considered in the analysis. The measurement units used throughout this study are rand denominated. The foreign cash, equity and bonds, gold index and the commodities index were converted to rand using an average rand/dollar exchange rate at the particular month.

The following steps will be included in our analysis:

1. Preliminary analysis of the properties and clustering structure of the Economic data chosen.

2. Perform a PCA decomposition of the economic data for dimension reduction.

3. Use the reduced data to perform a Fuzzy Cluster Analysis for different economic regime specification.

4. Evaluate and analyse the performance of the individual asset classes relative to specified benchmarks under the different regimes.

5. Perform an asset allocation optimisation for each regime and assess the performance of such a strategy to traditional asset allocation strategies that do not consider regime shifts.

The analysis that follows in this thesis will be done using the R – statistical software. Further acknowledgement is made to Balkissoon,

Bennett, Boudt, Carl, Leste, Ulrich, Wuertz and Zivot (2014) for Performance Analytics and Portfolio Analytics package in R that was used extensively in this thesis.

2.3 SCOPE AND DELIMITATIONS

This study was delimited to the analysis of an "economic regime cognizant" asset allocation strategy. In so doing, it was not in the interest or focus of this study to test the robastness of the regime classification method but rather propose a method that will be used for illustrative purposes. Thus, no robustness test was performed on the proposed models that are used to achieve the main objectives of this study.

As with Blitz and van Vliet (2004), the biggest limitation of this study was a lack of an out-of-sample testing opportunity. This was due to limited data availabile for such an exercise. That is, in order to be able to perform an out-of-sample test, one would have required a test data that would have encompassed all the defined regimes. In our data, an average regime covered 18 - 20 months. Given that this study assessed 4 regimes, we would have needed at least 6 to 7 years of data for an out of sample test.

This requirement could not be met given our limited data set as will be explained in Chapter 3 of this study. The study of Blitz and van Vleit (2004) was also conducted under the same constraints as this study where the authors used 20 years of data. Blitz and van Vleit (2004) highlighted their inability to perform an out-of-sample test and argued that even the 20 year in-sample data was not sufficient for the purpose of regime classification.

Thus, in this study we do not perform an out of sample test. However, we provide an analysis of an in-sample performance test which motivates for an out of sample testing for furtherwork. We argue that it is most important to perform the in-sample test since if a strategy fails to perform under controlled conditions with 100% foresight of the investment environment, it should not warrant out-of-sample testing. However, a strategy that shows attractive performance attributes in an in-sample test should warrant great support and reason for an out-of-sample testing hence the need of an in-sample testing.

Part II

DATA AND METHODOLOGY

3

DATA

3.1 ECONOMIC DATA

This section introduces the variables that are used in the process of economic regime classification. The data will consist of 8 macroeconomic variables and 5 financial factors. The aim behind using a combination of economic and financial variable is to establish a set of factors that can be used to best describe the investment environment. Table 2 below provides the variables and their publication frequencies. All the data was collected from Thomson Reuters Datastream Professional.

3.1.1 *Choice of the study period*

The data covers a period of 21 years starting from June 1993 to December 2014. The reason for not using a longer data set was motivated by the major structural change that occurred in the financial and economic environment of South Africa prior 1993. During 1990 – 1993, the 1985 trade and fiscal sanctions that were imposed on South Africa were being removed (Levy, 1999). This transitionary period is one which could not be expected to reoccur hence renders the market structure and the macroeconomic environment to be unique. That is, the behaviour of the market structure and the economy were driven by factors that would not be persistant and were not likely to reoccur in the future.

Since these trade sanctions were introduced in early 1985, this makes the period 1985 – 1993 not a good representative of the correlation structure between the market and the real economy. Furthermore, the episode of the financial rand introduced in the early 1960's further

Table 2: Set of variables used for Regime Classification

Variable	Frequency		Variable	Frequency
GDP	Quarterly		Credit Spreads	Quarterly
Inflation	Monthly		Earnings Yields	Monthly
Productivity	Quarterly		Dividend Yields	Monthly
Man. (Order & Sales)	Monthly		M1 Money Supply	Monthly
Current Acc: GDP Ratio	Quarterly		Yield Curve Level	Monthly
Real Trade Weighted Index	Monthly		Retail Sale	Monthly

separates the current and future market behaviour from that which persisted during 1960 – 1985 (Eun, Kilc and Lai, 2012). Thus, the relationship between the macro economy and the financial market prior 1993 was not stable and consistent as the period post 1993. Hence the inclusion of this period in our analysis would distort correlation behaviour and relationships.

This however does not imply that the period post 1993 did not experience significant structural changes. There are two policy reforms that are worth mentioning that impacted both the macro–environment and the financial market environment. The first amongst these is the introduction of the Growth, Employment and Redistribution (GEAR) policy in 1996 (Lewis, 2001). The purpose of this policy was to reduce the level of inflation and budget deficit in the country while maintaining satisfactory levels of economic growth and development. During the years after this announcement, there was a significant improvement in inflation which resulted in reduced real cost of capital and increased investment within the private sector. This policy in effect affected the levels and correlation between major economic variables and financial markets. Secondly, the inflation-targeting monetary policy framework introduced in 2000 reduced inflation risks associated with the country. This instilled credibility of the South African Reserve Bank and increased business activities in key economic sectors of the country (van der Merwe, 2004). This also provided a more structural and defined relationship between the macro environment and the financial market.

3.1.2 *Variable Selection*

The choice of economic variables used in this study was largely influenced by the works of Han and Turner (2010) and availability of data for each series. Han and Turner (2010) used a set of 29 economic variables in which a PCA decomposition was applied. However, we highlight a few other motivations for the variables used in this study.

Firstly, in most academic literature, the most common variables used to gain an understanding of a countries economic condition are Inflation and GDP. This includes the works of Kollar (2003) and Eychenne and Martinetti (2011). Similarly, in practice, most research analyst use these two variables to gauge the regime in which the economy is in (Munro and Silberman, 2008). Furthermore, the construction of the business cycle rests heavily on the behaviour of GDP. The South African Reserve considers flactuation around the long run trend of GDP as a way of measuring the business cycle (van der Merwe, 2004).

Vleit and Blitz (2009) extended the bi-variable framework (GDP and Inflation) for regime construction to a quadric-variable framework. In their work, the authors introduced the unemployment rate

and manufacturing order and sales as macro variables together with two financial factors (earnings yield and credit spreads). The aim of this extended framework was to show how the use of more variables to model the investment environment improves the approximation.

The money supply (M1) and the level of the yield curve as measured by the difference between 10 year yields and 3 month treasuries are the only variables that are included in this study but no literature was found that uses them. These variables were chosen subjectively through consultation with industrial practitioners and for their presence in the computation of the leading market indicator for the South African economy.

3.1.3 *Data Structuring and Adjustment*

The analysis in this thesis is done using monthly data. In order to establish monthly data for the time series that have quarterly data, we used a cubic-spline interpolator. Furthermore, we use seasonally adjusted inflation rates, real GDP figures and the real trade-weighted currency index. The quarterly data are lagged according to the time it takes to receive them in order to avoid look-ahead bias.

Drawing from the work of Han and Turner (2010), the economic condition at each point in time in the period from 1993 – 2014 can be explained using the economic and financial data collected. In order to reduce the complexity of the analysis and the noise embedded in other variables, we perform a Principal Component Analysis (PCA) on the standardised data series. In so doing, we make the assumption that each economic state can be presented as a projection to the first "n" PC's where n represents the number of chosen chosen principal components. Similar to Han and Turner (2010), we use the first "n" PC's that explain at least 80% of the variation in the data set as a selection criteria. As will be demonstrated in a chapter 4, the main aim of the principal component decomposition is to ensure that, time points of similar characteristics cluster together using a reduced dimensional space.

In order to ensure that the clustering of similar periods in the reduced data space occurs meaningfully, an assessment is required to establish whether the decomposed results make sense fundamentally. In order to verify this, we provide a PCA biplot relationship that projects the sample data points to the first two principal components. The representation of the PCA biplot is given in Figure 3 on the next page.

Figure 3 provides a biplot relationship between the first and second principal component. In order to obtain the chart, the bi-plot was performed on the time periods. This representation is different to that of Han and Turner (2004) given in figure Figure 1 in the literature review. That is, Han and Turner (2004) used dots rather than the actual time

Biplot Relationship of 1st and 2nd PC
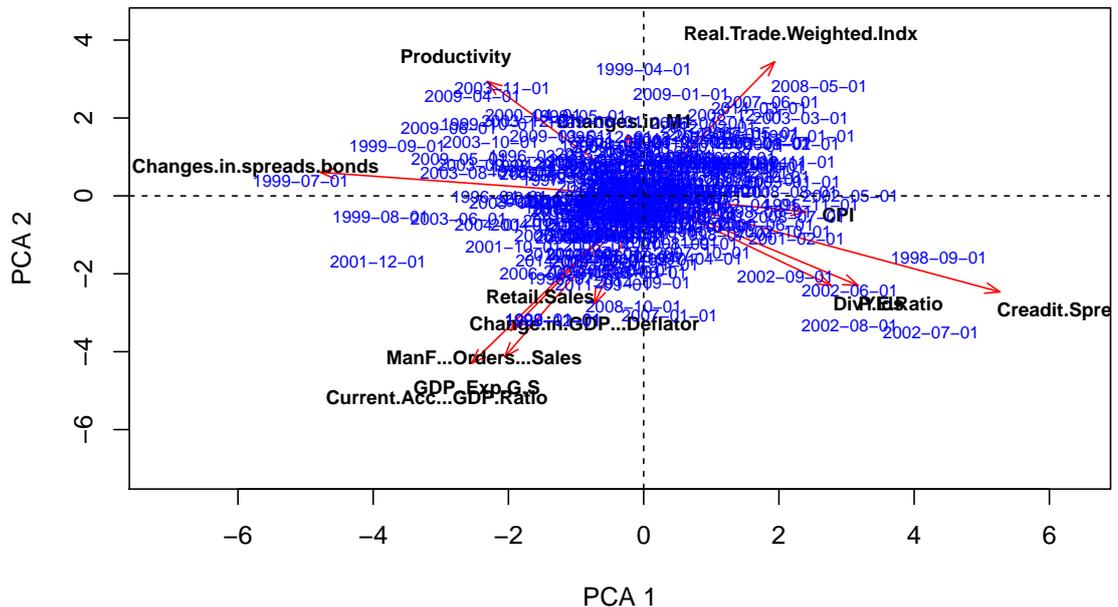
Figure 3: Biplot analysis of variables

period. Although Figure 3 may appear to be chaotic at first glance due to the chosen style of presentation, it does portray a vivid image of the relationship between the time periods and the variables. Relative to the first PC, the bottom left of the plot can be associated with positive economic indicators. That is, GDP, retail order and sales, manufacturing orders and sales etc. are all positive indicators where an increase in them implies a positive economic out-look. On the other hand, the top half with respect to PC 2 is associated with the cost of capital in the economy that may influence economic activities. That is, changes in bond spreads, productivity and the trade-weighted index are highly sensitive to changes in the cost of capital. The bottom left quadrant is associated with the inflation matrix that drives the investment environment.

The clustering of the time periods in the data set provides reasonable comfort of a sound and sensible clustering of the data. We further investigate if we can identify major turbulences that occurred within the analysis period from the PC's. Since the first PC accounts for the most variation in the data, we would expect that it accounts for periods of major economic shocks. Figure 4 on page 23 provides a projection of the state of the economy over the period to the first PC.

It is evident from Figure 4 that the first PC provides the relevant information of the different turbulent periods that occurred in the period under study. That is, recessionary periods and some minor structural disturbances in the South African economy can be deduced from the behaviour of the first PC. Furthermore, from Figure 4, we

Projection of Economic State into the First PC
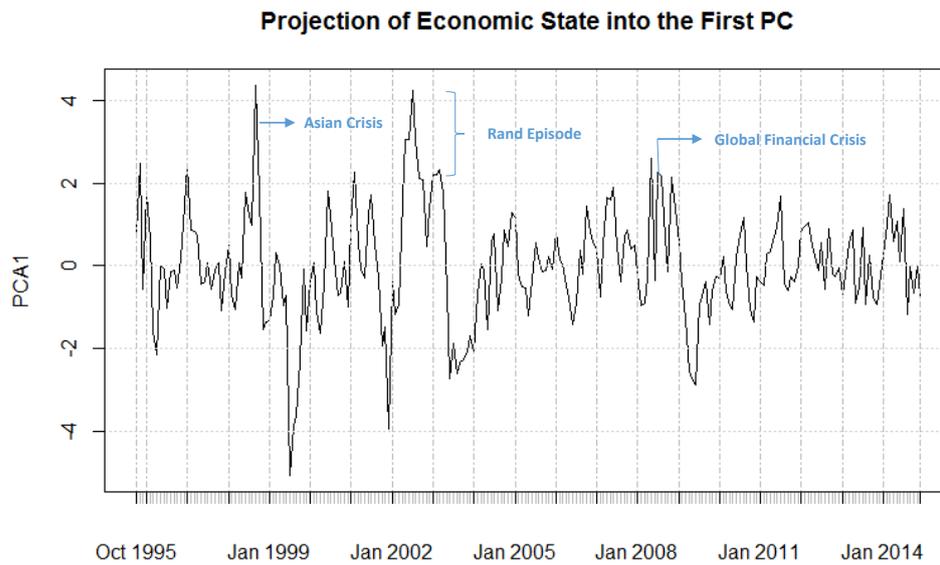
Figure 4: Relationship between the first PC and economic events

can see that major spikes in the first PC are associated with turbulent conditions such as the 1998 – 1999 emerging market crises, the 2001-2002 South African currency crises and the 2008 – 2009 financial crisis. The above analysis provides a validation on the use of the reduced data in order to model the different regimes the economy transitioned through over the period. This certifies that the data is in a sensible and stable structure that can be used to model the different economic regimes through the cluster analysis method applied on the reduced data.

3.2 ASSET CLASS DATA

In this study, we consider 10 asset classes for the analysis. We use the monthly return series for each of the asset classes for the period from 1993 to 2014. Table 3 at the top of page 24 provides the names of the asset classes used, source codes and proxies used in extracting the return series from I-Net Bridge.

Foreign cash, foreign bonds, foreign equity, gold, and commodities were all initially denominated in dollars hence we converted these to rands. The method of conversion from dollar to rands utilized the average monthly rand/dollar exchange rate for each month.

Table 3: List of Asset Classes Considered

Asset Classes	Code	Proxy
Local Cash	TBT3	3-Month Treasury Bills
Local Bonds	JAPI05	All Bond Index (ALBI)
Local Equity	J203T	All Share Index (ALSI)
Inflation Linked Bonds	BSAGI	Barclays SA Govt. ILB Index
Property	J253T	SA Property Index
Foreign Cash	USTB3M	US-3-Month Treasuries
Foreign Bonds	GLOUS	JPM Global Bond Index
Foreign Equity	MSCI.WORLD	Global Equity Index
Gold	GOLR	Global Gold Index
Commodities	FCRB	Commodity Index

3.2.1 *Asset Class Properties*

Figure 5 at the top of page 26 provides the cumulative returns for the different asset classes from the period 1993 to 2014. It is quite clear from Figure 5 that no single asset class exhibits cumulative returns that are consistently higher than the other asset classes. Initially, we observe that local equity outperformed the other asset classes from mid 1993 till early 1995 on a cumulative return bases. This outperformance was later substituted by foreign equity and at the later stage by local property.

Clearly from Figure 5 an investor would have been better off if they had a relatively high allocation to the property asset class. However, such a portfolio would have experienced major drawdowns in the period 1998 − 1999 and 2007-2009 period. Furthermore, portfolio managers are constrained by certain mandates which might have restrictions on the weight allocation on the different asset classes.

3.2.2 *Correlation Structure of Asset Classes*

Correlation amongst asset classes is a very important measure that is used in many asset allocation strategies to ensure diversification in the portfolio. Chin (2013) demonstrated in his work that these correlations between asset classes tend to have different "personalities" in behaviour over time. Table 4 on the next page provides the correlation measures amongst the asset classes in this study over the entire study period.

In Table 4, we observe that over the period, there appears to be a positive relationship between local equity and local bonds, gold and commodities while a negative correlation exist between local equity

Table 4: Correlation Matrix of the Asset Classes over the period 1993 to 2014

	Cash	Bonds	Equity	ILB	Property	F. Cash	F. Bonds	F. Equity	Gold	Commodities
Cash	1	0.23	-0.48	0.45	0.04	0.47	0.70	0.26	-0.07	0.32
Bonds	0.23	1	0.34	0.41	-0.14	-0.10	0.03	-0.09	-0.05	-0.11
Equity	-0.48	0.34	1	0.03	0.28	-0.09	-0.21	0.07	0.51	0.23
ILB	0.45	0.41	0.03	1	-0.14	0.18	0.25	0.31	0.30	0.34
Property	0.04	-0.14	0.28	-0.14	1	0.24	0.28	0.29	0.51	0.56
F. Cash	0.47	-0.10	-0.09	0.18	0.24	1	0.78	0.27	0.35	0.65
F. Bonds	0.70	0.03	-0.21	0.25	0.28	0.78	1	0.63	-0.05	0.45
F. Equity	0.26	-0.09	0.07	0.31	0.29	0.27	0.63	1	-0.02	0.32
Gold	-0.07	-0.05	0.51	0.30	0.51	0.35	-0.05	-0.02	1	0.80
Commodities	0.32	-0.11	0.23	0.34	0.56	0.65	0.45	0.32	0.80	1

Cumulative returns of Asset Classes
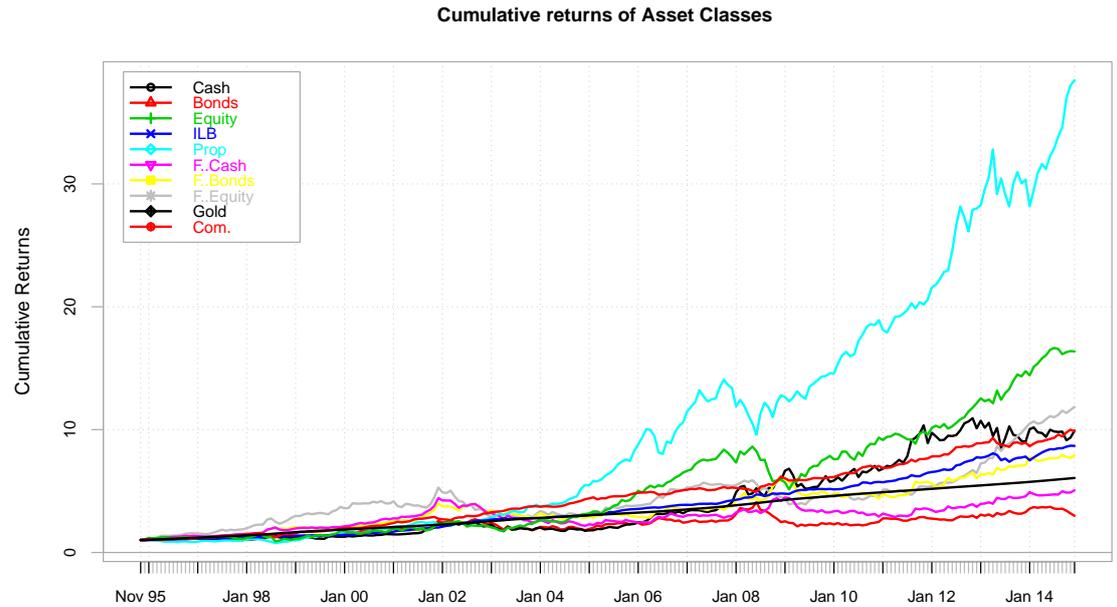
Figure 5: Asset Class Cumulative returns over the study period

and foreign bonds. Local bonds showed a negative correlation with foreign equity, gold and commodities while it had a positive correlation with foreign bonds. Chin (2013) points out however that correlations amongst asset that are time-period dependent. This implies that optimisation processes that use correlation matrices make the assumption that the choice of a time period is structurally stable and consistent.

In order to gain more understanding on the behaviour of the correlation structure amongst the different assets, we consider a monthly rolling period with a window of 3-years. We use a window of 3-years subjectively based on other studies in the literature that have used a length of 3-years. Figure 6 on page 27 provides the rolling correlations of a selected pairs of asset classes.

Figure 6 demonstrates the cyclical nature embedded in correlation structures amongst asset classes. We observe that the correlation between equity vs property, cash vs property and cash vs equity demonstrates periods of positive and periods of negative correlation over the period. This is even more evident amongst equity and bonds where the move from positive to negative and back to positive correlations has occurred more frequently.

This behaviour is consistent with that found by Chin (2013) in his study. This behaviour has been shown in many other academic literature where the focus was on how this result can be utilised to better gauge diversification within portfolios. In this study however, we use this result as a supportive argument on the differential in behaviour
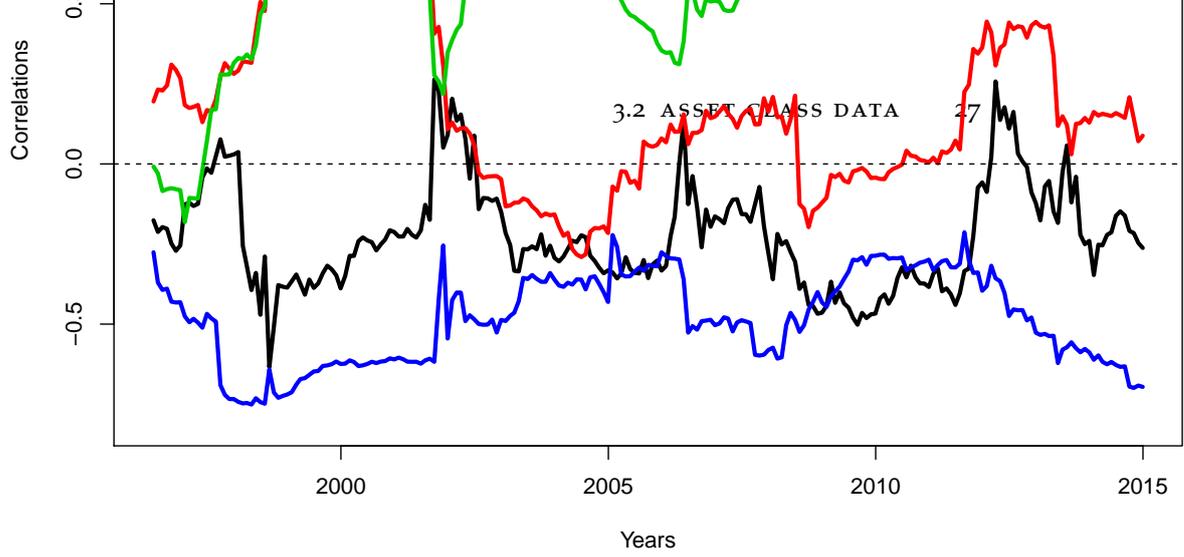
Figure 6: Rolling Correlation of asset classes

of asset classes across different economic regimes. This is mainly due to a similar argument presented by Chin (2013) that these correlation cycles across time reflect changes amongst asset classes in response to different market cycles and market events.

METHODOLOGY

As initialy outlined in the thesis statement, the aim of this dissertation is to develop an asset allocation strategy that takes account of cyclical changes in the investment environment while providing an attractive risk return profile. Drawing from the works of van Vleit and Blitz (2009), this chapter provide the methodology that will be followed for the proposed regime-based asset allocation strategy that seeks to accomplish this aim.

The portfolio construction process provided in this thesis is divided into three parts. Firstly, the different regimes that the economy undergoes during the period under study is defined. Secondly, we investigate asset class performance under each regime and find an optimal asset allocation for each regime. Lastly, we assess the performance of the resultant allocation strategy relative to common investment benchmarks to establish any significant outperformance.

The remainder of this chapter will be devoted to outlining the methods followed in each of the parts mentiond. Figure 7 on page 30 provides a flow diagram of the process that will be followed.

4.1 ECONOMIC REGIME CLASSIFICATION

To the best of our knowledge, there is no standard method suggested in academic litreture that classifies time periods into economic regimes for asset allocation purposes. The majority of literature reviewed in chapter 2 on this regard pointed out the subjectivity involved in justifing the defined regimes. For our purpose, we initiate the regime identification process in a similar manner as Han and Turner (2010) through Principal Component Analysis (PCA) on the centred data of economic variables.

However, contrary to Han and Turner (2010) - where the authors used k-means clustering and Markov Switch model - we propose the use of a fuzzy c-means clustering approach. In this case, the clustering is done on pre-defined number of clusters, c, where each cluster will be representing an economic regime.

We proceed by providing a brief and informal background on the theory of Principal Component Analysis (PCA) and Cluster Analysis with reference to how it will be implemented in the context of regime classification.
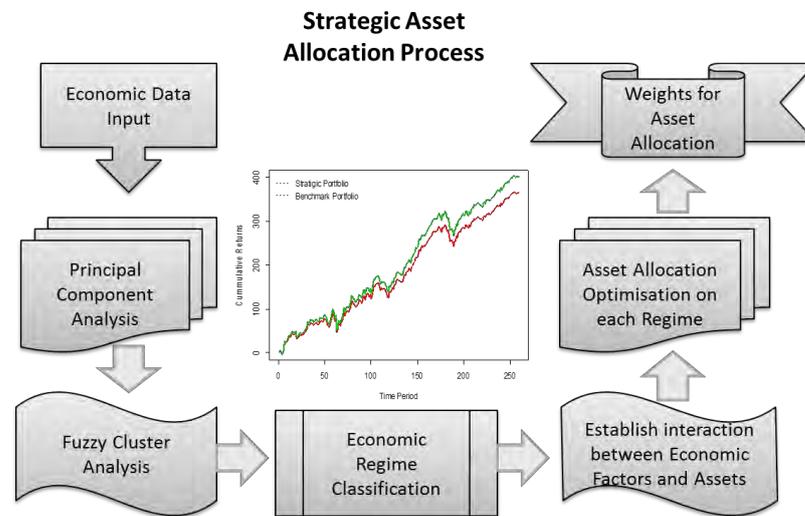
Figure 7: Strategic Asset Allocation Process

4.1.1 *Theoretical Framework of PCA*

Principal Component Analysis (PCA) is a multivariate statistical procedure concerned with dimension reduction through constructing a set of statistical factors that optimally captures major sources of variance within a data set (Jolliffe, 2002). These statistical factors, in terms of weighting schemes, provide information that can be used to explain co-movements in the original data which would have otherwise been complex to analyse. That is, instead of working with a matrix with dimensions $n \times m$, we can work with a new reduced matrix of dimensions $n \times p$ where $p < m$ while retaining most of the information of the initial matrix (Jolliffe, 2002).

This is particularly useful in our case where the aim is to provide a more realistic approach to regime classification. That is, contrary to the studies that use 2 or at most four economic variables to construct their regimes, a broader set of economic variables as descriptive of the state of the economy is considered.

The objective of the PCA procedure is finding a transformation matrix that transforms the original data matrix such that the transformed data optimally captures the variability exhibited in the original data while minimising redundancies (Jolliffe, 2002). The computed statis-

tical factors populate the columns of the transformation matrix and are termed principal components.

The PCA procedure is based on a number of assumptions. Amongst these assumptions, we assume the relationship between the variables of the original data can be expressed by a linear relation (Jolliffe, 2002). This assumption has two significant implications. Firstly, the assumption of linearity provides an easy way to restrict the potential set of basis matrixes that can be used for transformation (Jolliffe, 2002). Secondly, each of these statistical factors computed to re-express the original data becomes a linear combination of the original variables. This formalises the implied assumption of continuity within the data set.

Furthermore, when computing these statistical factors, or rather the principal components, the procedure used ensures that they are orthogonal to each other (Jolliffe, 2002). The need for orthogonality stems from the requirement of the transformation matrix to optimally capture variation in the original data without redundancies. Thus orthogonality implies that the covariance matrix of the transformed data is diagonal with zero entries on the off-diagonals hence zero correlations which reduces redundancies. Although the PCA procedure can be directly applied to the original data, when measurement units of the variables differ, as it is in our case, the covariance (or correlation matrix) of the standadised data is used.

The PCA procedure establishes the aforementioned transformed matrix through eigenvalue decomposing of the covariance matrix of the original data. Using the same notation as Shlens (2003), we begin by defining the following variables:

X Matrix of the original data

$\tilde{X}$ Standardised original data

C_X Covariance matrix of X

P Transformation Matrix

Y Transformed Data

C_Y Covariance matrix of the transformed data

Thus, to formanlise the PCA objective, the aim is to transform the original data, X, to a new transformed data, Y, such that the covariance matrix, C_Y, is diagonalised and the transformation matrix P is orthonormal with respect to its columns (Shlens, 2003). Mathematically, we require a transformation matrix P, such that

$$XP = Y \tag{1}$$

where the covariance matrixes of X and Y are formulates as

$$C_X = \frac{1}{n-1}X'X \tag{2}$$

$$C_Y = \frac{1}{n-1}Y'Y \tag{3}$$

repectively.

As previously mentioned, in deriving the transformation matrix P, we apply an eigenvalue decomposition to the covariance matrix C_X such that C_Y is diagonalised. From equation 1 and 2, the relationship between the transformation matrix and the covariance matrix of the transformed data can be shown as

$$C_Y = \frac{1}{n-1}Y'Y = \frac{1}{n-1}(XP)'(XP) = \frac{1}{n-1}P'X'XP. \qquad (4)$$

It thus follows from equation 4 that the relationship between the covariance of the original data, C_X, and covariance of the transformed matrix, C_Y, can be shown as

$$C_Y = \frac{1}{n-1}P'X'XP = P'(\frac{1}{n-1}X'X)P = P'C_XP. \qquad (5)$$

Using results from linear algebra, Shlens (2005) showed that any symmetric matrix, say A, can be expressed as a product of matrixes of its eigenvalues and eigenvectors. That is, given a symetric matrix A, it can be re-express as

$$A = EDE' \qquad (6)$$

where E is a matrix of eigenvectors and D is a diagonal matrix of eigenvalues. Thus, since C_X is a symmetric, it can be decomposed in the same way as matrix A in equation 6. If we let $P = E$, then, following the result from equation 6, the covariance matrix C_X can be expressed as

$$C_X = \frac{1}{n-1}PDP' \qquad (7)$$

which implies that the covariance of the transformed matrix can then be expressed as

$$C_Y = P'C_XP = \frac{1}{n-1}P'(PDP')P = \frac{1}{n-1}(P'P)D(P'P). \qquad (8)$$

Since we have assumed an orthogonal matrix P, results from linear algebra show that $P' = P^{-1}$, thus equation 8 can be expressed as follow:

$$C_Y = \frac{1}{n-1}(P'P)D(PP') = \frac{1}{n-1}(P^{-1}P)D(P^{-1}P) = \frac{1}{n-1}D \quad (9)$$

which shows that the resultant matrix P from the eigenvalue decomposition diagonalises the covariance matrix C_Y which in effect accomplishes the desired goal for the PCA procedure.

The intuition behind the eigenvalue decomposition procedure is simple in theoretical terms and can be summarised in the following three steps (Shlens, 2003):

1. Given a data set, say X, the first principal component is obtained by finding a normalised direction in the m- dimensional space in which the variance in X is maximised. In order to capture this variance, the matrix C_X is used. This vector is stored as the first principal component p_1.

2. The second principal component, p_2, is then found by finding another directional vector along the m- dimentional space that maximises the variance in X that was not captured by the first component, p_1. Since we have assumed orthogonality in the principal components, the search for the second component is restricted to vectors orthogonal to the first component.

3. This process is repeated until we have m- such components where each p_i^{th}- component lies perpendicular to all other components, where $1 < i \leq m$.

The resulting matrix from the above process provides a set of principal components. The diagonalised matrix of eigenvalues is ordered in descending order, with the largest eigenvalue on the first column and the lowest eigenvalue on the last diagonal element of the eigenvalue matrix. The eigenvectors, or rather the principal components, are subsequently arranged in a similar manner.

The relative size of each diagonal element of the eigenvalue matrix expresses the proportion of variability of the original data as explaind by the coresponding principal component. Thus, it should easily follow that depending on the desired level of proximity of the reduced data, we can easily select the number of principal components that will match the requirements.

We implement the above eigen-decomposition on the macroeconomic data. As we mentioned, we apply the PCA procedure on the covariance matrix of the standardised matrix, $\tilde{X}$. That is, each column of X has a zero mean. The resultant principal components are then used as a representation of the evolution of the South African economy.

4.1.2 *Theoretical Framework of Cluster Analysis*

The Cluster Analysis framework is the backbone for the regime classification method followed in this thesis. The main objective of regime classification in this regard is to optimally identify cyclical phases in the period 1993 to 2014. In order to achieve this, we use a clustering method known as Fuzzy Cluster Analysis. In principle, clustering focuses on identifying time periods with similar characteristics (Holland, 2006). This principle becomes very useful for our purposes where we seek time periods of similar characteristics that would form economic regimes.

The main reason for using fuzzy clustering rather than hard clustering can be explained by considering the fundamental differences between these two methods. Unlike hard clustering, where we would be forced to classify each time period to a unique cluster from the outset, fuzzy clustering allows for time periods to partially belong to each cluster with different degrees of membership (Madhulatha, 2012).

This has two implications for regime classification. Firstly, fuzzy clustering allows for a more realistic approach of grouping time periods since we expect other time periods to exhibit characteristics of regimes that are not necessarily the dominant regime observed (Chad, and Giles, 2003). That is, we might be in an expansionary cycle in the economy but find two months of a negative print in GDP which resembles a recessionary period. Secondly, the resultant cluster centres will be more stable when we add new data which is a very attractive feature for forecasting purposes (Madhulatha, 2012). That is, when we add new data, we expect the cluster means to be stable such that we can calculate the distance to each cluster from an observation without worrying about the stability of the cluster centres.

Generally, clustering techniques are grouping methods based on the partitioning of data into distinct groups characterised by high within-group similarities and low between group similarities (Madhulatha, 2012). For quantitative data, the term "similarity" is defined mathematicaly using a distance norm. In this case, points closer to each other (in distance terms) are regarded as more similar relative to points further apart.

One of the merits of such techniques has been due to their independence to common statistical assumptions such as prior requirements of distributional properties which underpins most conventional statistical methods (Madhulatha, 2012). We begin by providing an overview of clustering methods followed by a detailed explanation of the fuzzy method and how we have used it for regime classification.

4.1.3   *Overview of Clustering Methods*

In the last decade, a large amount of literature has been devoted to the field of data mining. This has led to a number of different methods being suggested and formulated. Amongst these, Cluster Analysis has been given a great deal of attention with different methods devised to tackle different problems in different fields. In general, clustering methods can be divided into two broad categories. The first class of methods are defined as Hard (Exclusive or crispy) Clustering methods while the second is defined as the Fuzzy (Soft or Overlapping) Clustering methods (Madhulatha, 2012).

Hard Clustering has its roots on classical set theory. Under this clustering method, each object is forced to either belong or not belong to

a particular set or cluster. Thus each object will ultimately belong to one and only one cluster group amongst the set of clusters. The resultant clusters become mutually exclusive with each cluster having at least one object within it.

Fuzzy Clustering on the other hand is a relaxation of the non-overlapping nature of hard clustering. That is, in Fuzzy Clustering, objects are allowed to simultaneously belong to more than one cluster with different degree of membership. This provides a more natural way to partition objects since in some cases, object may exhibit characteristics of more than one group. This is particularly useful for grouping object on the boundaries of clusters. In this case, objects will not be forced to belong to a particular group but rather assigned in a proportional manner across groups depending on its degree of similarity to these groups. Furthermore, since hard clustering assumes a discrete partitioning approach, a fuzzy clustering approach relaxes this assumption and allows continuity which permits the use of algorithms based on analytic functions for clustering.

4.1.4 *Basic Notations*

For completeness, we provide detailed explanation of both hard and fuzzy clustering. However, before we delve into the formulation of the methods, we begin by providing some basic notation for data representation. We define the input data from the PCA decomposition explained in *section* 1.1.1 as Z. That is,

$$Z = \hat{P}$$

where $\hat{P}$ is a matrix containing the number of chosen principal components to be used for dimension reduction. If we assume that we selected q - principal components from P , then, the dimensions of the matrix Z are given as $n \times q$. The rows of the matrix $Z_{n \times q}$ are representative of the row measurements from the macro data. The columns on the other hand are the principal components of the matrix X . Thus the matrix Z can be represented as

$$Z_{n \times q} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1q} \\ z_{21} & z_{22} & \dots & z_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nq} \end{bmatrix} .$$

In the matrix $Z_{n \times q}$, we let the rows of the matrix be the objects to be grouped or clustered. This is synonymous to clustering the monthly periods from 1993 to 2014. In this case, we expect months of similar behaviour to cluster together relative to those that less resembles each other. For the remainder of this section, we refer to the matrix $Z_{n \times q}$ as given above.

4.1.5 *Hard Partitioning*

As initially mentioned, the goal of clustering is to partition a data set, in this case the matrix $Z_{n \times q}$, into c-clusters. For our purpose, we will make the assumption that the desired number of clusters needed is known before hand with acknowledgement that in other cases these clusters are not known prior. Thus, following the same notation as Bezdek et al (1983), we can define a hard clustering on the matrix $Z_{n \times q}$ as a family of subset $\{A_i \mid 1 \leq i \leq c\} \subset P(Z)$ which has the following properties:

$$\cup_{i=1}^c A_i = Z_{n \times q} \tag{10}$$

$$A_i \cap A_j = \phi, \quad 1 \leq i \neq j \leq c \tag{11}$$

$$\phi \subset A_i \subset Z_{n \times q}, \quad 1 \leq i \leq c. \tag{12}$$

Equation 10 to 12 expresses the definition of hard clustering in a mathematical form. That is, from equation 10 we have that the union of all subsets clusters denoted by A_i which fully utilising the objects in the matrix $Z_{n \times q}$. Secondly, from equation 11, we have that the interaction between the different clusters is an empty or null set while equation 12 shows that no single defined cluster should be an empty set nor contains all the data from Z. If we now define a partitioning matrix $U = [\mu_{ik}]_{c \times q}$ such that each row of the matrix U is a representative of the values from the membership function μ_i of the ith subset A_i of Z, then, it follows that the constituents of the matrix U are conditioned to the following constraints:

$$\mu_{ik} \in \{0,1\}, \quad 1 \leq i \leq c, 1 \leq k \leq q \tag{13}$$

$$\sum_{i=1}^c \mu_{ik} = 1, \quad 1 \leq k \leq q \tag{14}$$

$$0 < \sum_{k=1}^q \mu_{ik} < q, \quad 1 \leq i \leq c. \tag{15}$$

From equation 13 to 15, the following charecteristics of hard clastering can be deduced. Equation 13 provides a distinct atribute of hard clustering. That is, the constituents of the partitioning matrix U can only take on values $0 \, or \, 1$. In order to ensure that each row or object belongs into a cluster, equation 14 requires that the summation

for each object across the clusters should be one. Bezdek et al (1981) defined a space which contains all possible hard partitioning's as

$$M_{hc} = \left\{ U \in \mathbb{R}^{c \times q} \mid \mu_{ik} \in \{0,1\}, \forall i,k; \sum_{i=1}^{c} \mu_{ik} = 1, \forall k; 0 < \sum_{i=1}^{c} \mu_{ik} < q, \forall i \right\}.$$

$$(16)$$

The framework of hard partitioning provides a foundation in which the fuzzy method is built on. This is examined in the next section.

4.1.6 *Fuzzy Partitioning*

The Fuzzy partitioning method is a generalisation to the hard partitioning provided in subsection 1.2.3. From the hard partitioning method, the constraining condition given by equation 13 defines the crispiness of this method. That is, by letting each object μ_{ik} to assume either a 0 or 1 as given by equation 13 implies that each object can either belong (1) or not belong (0) to a group. It thus follows that in order to relax this hard form of clustering, we need to allow each object to assume any real value in the interval $[0,1]$. Hence, analogues to hard clustering, conditions for the Fuzzy Clustering can be given as:

$$\mu_{ik} \in [0,1] \tag{17}$$

$$\sum_{i=1}^{c} \mu_{ik} = 1, \quad 1 \le k \le q \tag{18}$$

$$0 < \sum_{k=1}^{q} \mu_{ik} < q, \quad 1 \le i \le c. \tag{19}$$

Thus, from equation 17 to 19, we have that the partitioning matrix U has elements that are allocated in a fuzzy manner with the requirement that the sum of the proportional allocation of any object across the clusters to be 1. In this case, the total membership of each object in $Z_{n \times q}$ will equal to 1. This form of partitioning or rather allowing the data to cluster in such a manner is an attractive feature for the purpose of regime classification. To elaborate on this point, we make the assumption that no given period is likely to strictly exhibit characteristic of a single regime. Thus the fuzziness allows us to form realistic regimes which as will be shown later to provides an attractive scenario building and forecasting property. In the case of Fuzzy Clustering, Bezdek et al (1981) defined a space which contains all possible hard partitioning's as:

$$M_{fc} = \left\{ U \in \mathbb{R}^{c \times q} \mid \mu_{ik} \in [0,1], \forall i,k; \sum_{i=1}^{c} \mu_{ik} = 1, \forall k; 0 < \sum_{i=1}^{c} \mu_{ik} < q, \forall i \right\}.$$

$$(20)$$

4.1.7 *Fuzzy c-Means Clustering*

Amongst the fuzzy clustering algorithms available, we use the Fuzzy *c*- Means Clustering procedure in order to establish the regimes. This class of Fuzzy partitioning method is based on optimising a *c*- means objective function with pre-defined cluster numbers. The basic form of the c-means function is given as

$$J(Z, U, V) = \sum_{i=1}^{c} \sum_{k=1}^{q} (\mu_{ik})^m \parallel z_k - v_i \parallel_A^2 \tag{21}$$

where

$$U = [\mu_{ik}] \in M_{fc} \tag{22}$$

as initially defined and

$$V = [v_1, v_2, \ldots, v_c], \quad v_i \in \mathbb{R}^n \tag{23}$$

is a vector who's elements are the different cluster centres which are determined through the algorithm,

$$D_{ikA}^2 = \parallel z_k - v_i \parallel_A^2 = (z_k - v_i)^T A (z_k - v_i) \tag{24}$$

is a squared inner-product distance norm used for the similarities measure and

$$m \in [0, \infty) \tag{25}$$

is the fuzziness parameter that varies the level of fuzziness in the clustering process. That is, the value of m = 1 is equivalent to a normal hard clustering approach while as the value of $m \to \infty$, the clustering become completely fuzzy. The cost function value provides a measure of the total variance for each cluster around its centre. We use the fuzzy *c*- means function provided in this section for regime clasification where we have a pre-specified value for *c*. In the next section, we formalise the process of regime classification by providing a formalised regime classification algorithm that was followed for regime classification.

4.2   REGIME CLASSIFICATION

Given the theoretical background provided in section 1.1 and 1.2, we provides a formalised process which we follow for regime classification using these theoretical concepts outlined. Considering equation 24 above, equation 21 can be restated as

$$J(Z, U, V) = \sum_{i=1}^{c} \sum_{k=1}^{q} (\mu_{ik})^m D_{ikA}^2. \tag{26}$$

Minimisation of equation 26 is equivalent to finding stationery points of this function. Given the constrains of the problem (equation 17 to 19), Bezdek et al (1981) proposed the use of the Lagrangian method where the Lagrange multiplier is given as

$$\bar{J}(D, U, V, \lambda) = \sum_{i=1}^{c} \sum_{k=1}^{q} (\mu_{ik})^m D_{ikA}^2 + \sum_{k=1}^{q} \lambda_k \left[\sum_{i=1}^{c} \mu_{ik} - 1 \right]. \quad (27)$$

In the case where $D_{ikA}^2 > 0 \, \forall i, k$ and $m > 1$, Bezdek et al (1981) argued that $(U, V) \in M_{fc}$ will minimise equation 21 if

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} (D_{ikA}/D_{jkA})^{2/(m-1)}}, \quad 1 \leq i \leq c, 1 \leq k \leq q \quad (28)$$

and

$$v_i = \frac{\sum_{k=1}^{q} (\mu_{ik})^m z_k}{\sum_{k=1}^{q} (\mu_{ik})^m}; \quad 1 \leq i \leq c. \quad (29)$$

Equation 29 provides a set of first order conditions necessary for the stationarity of the function in equation 26. The clustering algorithm iterates through equation 28 and 29 in order to establish the necessary regimes. Bezdek (1980) provides a formalised proof of the convergence of the above conditions.

4.3   REGIME CLASSIFICATION ALGORITHM

Given the theoretical background provided in section 1.1 and 1.2 above, the fuzzy c- means regime classification algorithim can be given as

$$Algorithm\,1:\ Fuzzy\,c-Means\,(FCM)$$

1. Begin by specifying the following parameters:

   - $c\ \rightarrow\ $ The number of clusters (for our perposes, we let this be 4)

   - $m\rightarrow$ The fuzzyness parameter (we obtain this using cross validation)

   - $\epsilon\ \ \rightarrow$ The termination parameter (for our purposes, we use 0.001

   - $A\ \ \ \rightarrow$ A norm-inducing matrix of which we use the Euclidean norm defined in equation 24

2. Initialise the partitioning matrix $U^{(0)} \in M_{fc}$

3. For $\{\ell$ in 1 to ...$\}$

   a) Step 1: Compute the different clustering means:

   $$v_i = \frac{\sum_{k=1}^{q}\left(\mu_{ik}^{(\ell-1)}\right)^m z_k}{\sum_{k=1}^{q}(\mu_{ik}^{(\ell-1)})^m};\quad 1 \le i \le c.$$

   b) Step 2: Compute the distances:

   $$D_{ikA}^2 = (z_k - v_i^\ell)^T A(z_k - v_i^\ell),\quad 1 \le i \le c,\, 1 \le k \le q$$

   c) Step 3: Update the partition matrix:
      for $\{1 \le k \le q\}$ if $\{D_{ikA} > 0,\, \forall i\}$ then

   $$\mu_{ik}^{(\ell)} = \frac{1}{\sum_{j=1}^{c}(D_{ikA}/D_{jkA})^{2/(m-1)}}$$

   otherwise

   $$\mu_{ik}^{(\ell)} = 0\ if\ D_{ik}^2 > 0,\ and\ \mu_{ik}^{(\ell)} \in [0,1]\ with\ \sum_{i=1}^{c}\mu_{ik}^{(\ell)} = 1$$

4. Repeat the steps in point 3 until $\|U^{(\ell)} - U^{(\ell-1)}\| < \epsilon$.

The Fuzzy c-means algorithm will tend to produce different results for different starting or initialisation points. Hence, the algorithm con-

verges to local minima's for the c-means function. Secondly, Bezdek (1980) noted that the optimisation process as given above loops over the estimates $U^{(\ell-1)} \to V^{(\ell)} \to U^{(\ell)}$ with termination occuring when $\|U^{(\ell)} - U^{(\ell-1)}\| < \epsilon$. However, the optimisation can also be done by first initialising $V^{(0)}$ and looping through $V^{(\ell-1)} \to U^{(\ell)} \to V^{(\ell)}$ with termination criteria given by $\|V^{(\ell)} - V^{(\ell-1)}\| < \epsilon$. For our purpose, we use the former method with the resulting partitioning matrix U providing classification of each time point to a cluster with some degree of membership.

4.3.1 *Subjective Adjustment*

The nature of fuzzy clustering allows each object to belong to more than one cluster with different degrees of membership. When classifying a particular point with partial membership on different clusters, we subjectively consider the cluster where the point has the highest degree of membership. In the case where the fuzziness was equally proportioned (or close to equality), we considered the correlation of the behaviour of variables with the principal components at that particular time period. We subsequently use the most correlated variable to define the prevailing regime.

4.4 STRATEGIC PORTFOLIO SELECTION AND OPTIMISATION

For the asset allocation and optimisation process, we use the centroid approach initially suggested by Almgren and Chriss (2004). We provide a brief background on how the asset allocation process is carried together with the selection of an optimal portfolio.

4.4.1 *Asset Allocation Process*

The asset allocation method proposed by Almgren and Chriss (2004) is based on the notion of using ordinal information rather than quantitative estimates of return for asset selection. For asset selection, Almgren and Chriss (2004) assumed that the attractiveness of an assets is governed by the ordinal ranking of its expected return vector relative to other asset that is consistent with investors view. This use of ordinal information in investment theory and practice is well documented in academic literature. This includes the works of Fama and French (1992), Banz (1981), Chan and Lakonishock (2004), etc (Almgren and Chriss, 2004).

Academic scholars have demonstrated how the use of sorts based on the analysis of correlation between asset numerical factors (firm characteristics and price history) and expected returns can improve performance. This fact becomes even more prominent and attractive when these observed correlations are expected to persist going for-

ward. In this case, the ordinal structure defined by the sorting signature of expected performances can also be assumed to persist going forward.

Thus, from the above mentioned, we can deduce that the use of asset return ranking for allocation purposes is warranted if and only if the fundamentals on which the ranking criteria is founded upon can be assumed to persist within the investment environment. This condition provides great support for regime based asset allocation as used in this thesis. This is due to the idea of using regimes as a means of exploiting the cyclical nature embedded in assets within the investment environment. That is, we expect correlations of asset performance and regimes to persist on an ongoing basis given that the regimes have been well defined.

4.4.1.1 *Mathematical Formulation*

If we consider an investment universe of n stocks, then, Algrem and Chriss (2004) defined a sorting signature of a complete sort as a set of inequality relations between expected returns such that we have

$$r_1 \geq r_2 \geq \ldots r_n \tag{30}$$

where r_i is the expected return of the ith asset. A number of sorting signatures are provided by the authors. This includes Index sorts, Higher Order sorts, and Multiple sorts. We will restrict our analysis to using a complete single sort as given in equation 30.

Then, following from equation 30, the allocation between assets is further governed by a preference relation that is assumed to dictate investor rational. Mathematically, the preference relation between two portfolio's, $w \& v$, as defined by Almgren and Chriss (2004) is expressed as

$$w - v = \sum_{i=1}^{n-1} \lambda_i e_i + \gamma \mathcal{O} \tag{31}$$

where an individual portfolio, w say, is given as

$$w = \sum_i \lambda_i e_i + \gamma \mathcal{O}. \tag{32}$$

Looking at the left-hand side of equation 31, λ_i is defined as any positive integer while $e_i = (0, \ldots, 0, 1, -1, 0, \ldots, 0)$ Is a fundamental portfolio that expresses positions assumed, that is , from e_i, we buy a one rand of the ith asset and sell one rand of asset $(i + 1)^{st}$ asset. Then the composition of portfolio w is preferred to that of portfolio v if each $\lambda_i \geq 0$ for every i in equation 31.

4.4.1.2 *Efficient Portfolios*

The asset allocation process as described by Almgren and Chriss (2004) is not based on the preference relation alone, but also on the

concept of portfolio efficiency. That is, if the set of portfolios that meet our preference relation is given by Γ, then, from the prefrence relation given by equation 31, Γ can be expressed mathematically as

$$\Gamma = \left\{ w \in \mathbb{R}^n \mid w \cdot V \cdot w \leq \sigma^2 \right\}. \tag{33}$$

Thus from equation 33, we have Γ representing the set of all asset allocations that meet our risk target or rather budget constraint. Then, a portfolio, say w in Γ is defined as efficient if there is no other portfolio $v \in \Gamma$ that is strictly preferable to w.

4.5   PORTFOLIO OPTIMISATION

The allocation process provided in section 1.4.1 yields a set of efficient portfolios with no single portfolio identified as optimal. Almgren and Chriss (2004) finalised their proposed portfolio selection approach by showing how to choose a single optimal portfolio from a set of efficient portfolios. That is, using the sorting signature and preference relation defined in section 1.4.1, Almgren and Chriss (2004) defined an optimal portfolio with respect to a sort as the *most preferable* under the preference relation. In order to find the *"most preferable"* portfolio amongst the set of efficient portfolio's, Almgren and Chriss (2004) defined a vector c as the centre of mass of the set space, Q, of consistent expected return.

 If we consider portfolio's $w \& v$ that we used previously, then the vector c is such that portfolio w is preferred to portfolio v if and only if the following relation holds:

$$w \cdot c \geq v \cdot c. \tag{34}$$

Almgren and Chriss (2004) defined the vector c as the centroid vector. Equation 34 implied that the preference relation could be characterised as a linear function where the *"most preferable"* portfolio is equivalently found by maximising the linear function c in the set Γ.

 Thus, in order to find an optimal portfolio, Almgren and Chriss (2004) proposed finding an optimal solution to the following linear programing problem with quadratic constraints:

$$max_w \ w \cdot c \tag{35}$$

subject to

$$w \cdot V \cdot w \leq \sigma^2. \tag{36}$$

The solution to the above optimisation problem is termed a centroid optimal solution. The novelty of this method lies in its natural ability to select a portfolio composition through the use of the centroid vector. To elaborate on this point, mathematically, for any type of sort,

there exist a corresponding centroid vector to match an optimal se-
lection. The authors suggested two ways of computing the centroid
vector. That is, through the use of Monte Carlo methods or through
the evaluation of integrals. For our purposes, we use analytical ap-
proximations as given by Alemgren and Chriss (2004). The suggested
computation consider a single sort of n assets and approximates the
j^{th} component of the centroid vector numerically as

$$c_{j,n} = N^{-1} \left[\frac{n + 1 - j - \alpha}{n - 2\alpha + 1} \right] \tag{37}$$

where $\alpha = A - Bn^{-\beta}$, $N^{-1}(\cdot)$ is the inverse cumulative normal distri-
bution. The value of $A = 0.4424$, $B = 0.1185$ and $\beta = 0.2$

Thus, the portfolio selection and optimisation that is performed in
each regime independently can be summarised by the following steps:

```
Step 1: Compute the centroid vector using equation 37
Step 2: Perform an optimisation of equation 35 with
constraints given by equation 36
Step 3: Extract the weights of the resultant portfolio and
let that represent the optimal allocation for the regime
```

We apply the above optimisation process on each regime resulting in
4 generic portfolios. For the optimisation process in step 2, we use
the statistical program R with recognition to the use of the Portfolio
Analytics Package by Balkissoon et al (2014).

Part III

RESULTS, DISCUSSIONS AND CONCLUSION

EMPIRICAL RESULTS

In this chapter the results obtained following the steps outlined in the methodology section are provided. This chapter begins by demonstrating how the different economic regimes were obtained from the cluster analysis process and how this method can be used to forecast forthcoming regimes. This will be followed by results on the behavior of the different asset classes within regimes.

5.1 IDENTIFICATION OF CLUSTERS

In order to define the different regimes, the classification process outlined in chapter 4 was followed. To recap, the process is initiated by a PCA decomposition on the centered macroeconomic data set. This is followed by a fuzzy cluster analysis on the matrix of a selective set of principal components. The actual clustering is done on the time periods (i.e. rows of the PCA matrix) and not the variables (principal components). In so doing, the goal is to group time periods with similar characteristics as captured by the behavior of the principal components.

It was also stated in chapter 4 that the number of desired clusters were pre-specified prior the clustering process. Four clusters were pre-specified with the aim of capturing the different regimes the economy cycled through. Figure 8 on the next page provides an illustration of the clusters by projecting them on the first two principal components. The reader is reminded that the actual clustering was done using the first 5 principal components and not the first two. Thus, Figure 8 provides the projection of clusters defined in a 5-dimentional space which have been projected into a 2-dimentional space. For ease of visualization, we have further ordered the months from June 1993 to December 2014 using the ordinal scale of 1 to 246. In this case, 1 on the cluster plot corresponds to June 1993 while 246 would correspond to December 2014.

The results depicted by Figure 8 are illustrative of the aim we hoped to achieve through the use of fuzzy clustering. It is evident that most of the time periods fall in more than one cluster. This is supportive of a hypothesis that seeks a more realistic approach in constructing regimes. That is, rather than forming clusters that classifies time periods using hard clustering, we allow the time periods to partially belong to clusters. Thus, the cluster means under the fuzzy clustering method are significantly different and - arguably - more
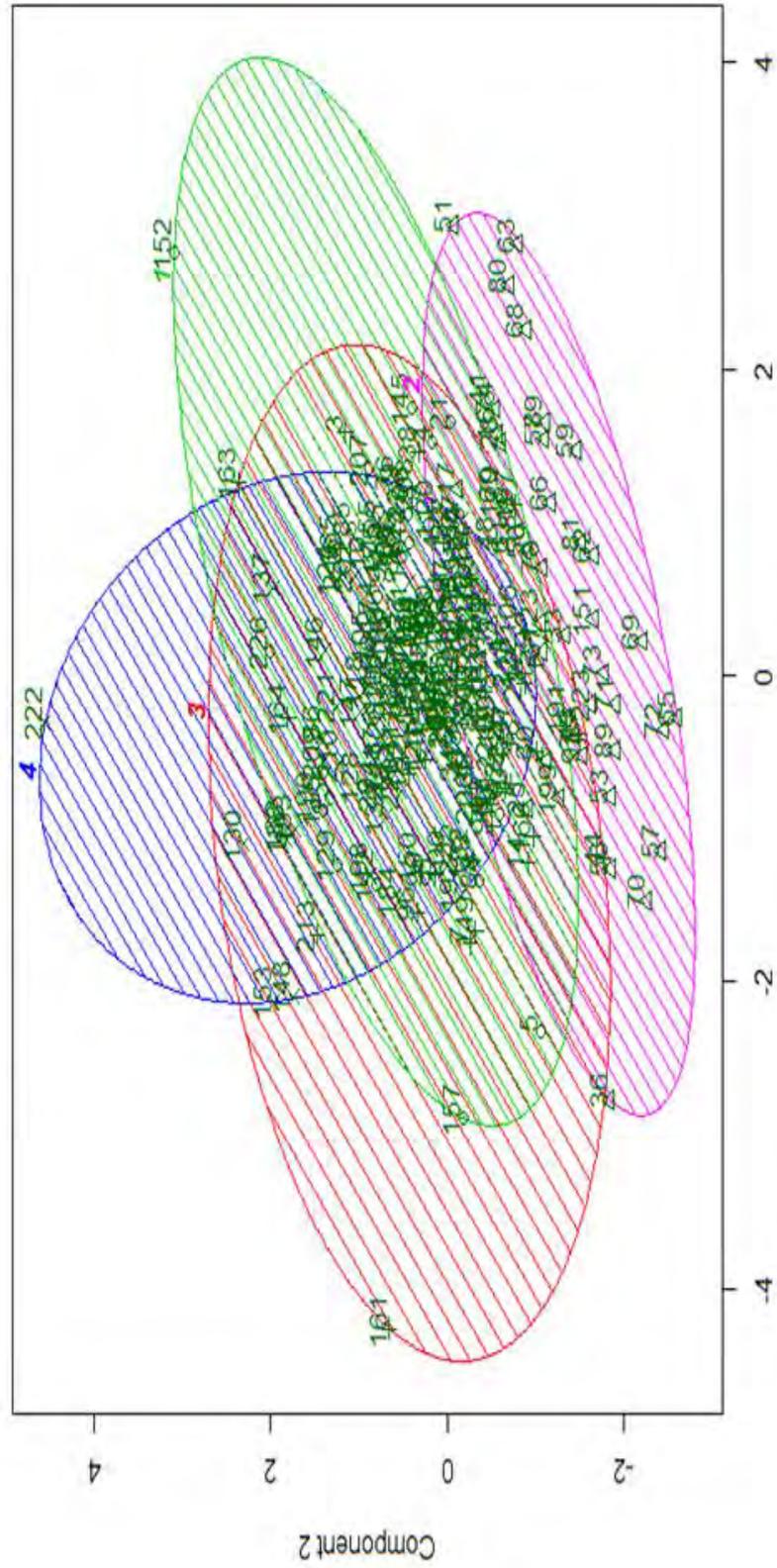
47

Figure 8: Projection of the resultant clusters to the first two principal components

Table 5: Behaviour of macro-variables across regimes

	Cluster Centers (Standardized Values)			
	Cluster1	Cluster 2	Cluster 3	Cluster 4
CPI	−0.39	−0.10	0.06	0.53
GDP	−0.36	−0.05	0.22	0.00
Productivity	0.04	1.82	−0.47	−0.47
Changes in GDP Deflator	−0.10	0.16	−0.19	0.36
Manf. Order and Sales	−0.02	0.08	−0.05	0.04
Retail Sales	−0.28	−0.11	0.14	0.12
Current Acc. to GDP	0.14	0.12	−0.03	−0.16
Real Trade Weighted Index	−0.04	0.06	0.05	−0.15
Credit Spreads	0.89	−0.16	0.23	0.20
PE Ratio	0.27	0.14	−0.24	0.09
Dividend Yields	−0.27	−0.08	0.16	0.00
Changes on M1 money supply	−0.13	−0.17	−0.08	0.42
Changes in Bond Spreads	0.66	−0.36	−0.13	−0.07

realistic in this context to those that would have otherwise resulted through the hard clustering process.

In order to translate the information portrayed by the clusters into meaningful regimes, we consider the behavior of the different economic variables within each cluster. This is done by first grouping all the time periods into their respective clusters. Since other time periods had different degrees of membership on the clusters, the ultimate classification process of any given time point (observation) was based on the smallest distance between the observation and the cluster means. Thus a time point that had characteristics of all four clusters would be grouped to a cluster with the smallest distance between the cluster mean and the observation. This allowed us to uniquely allocate each time period to a single cluster without changing the cluster means to be those of hard clustering.

After grouping the different time points into the different clusters, the original centred data was spliced according to the clusters. In so doing, each cluster contained data for each of the macro variables that corresponded to the time periods in the cluster. This was followed by calculating the mean value of each economic variable under each cluster. This is possible since the clustering was performed on the time points although the variables that were used were the principal components. The reason for reverting back to the original data is so that we can analyse and understand the behavior of the economic variables in each cluster and to subsequently define the different clus-

ters into regimes. Table 5 at the top of the privious page provides the means of the macro variable across the different clusters.

From Table 5, we view the means of the macro-variables as the implied cluster centers for the variables across each cluster. These variable means provide useful information on the characteristics of the different clusters and how we can translate these clusters into meaningful regimes. Clearly from Table 5, cluster 1 is associated with a decrease in most of the positive economic variables. That is, we observe that on average, CPI, GDP, manufacturing, dividend yields and retail sales are falling. This implies a negative economic outlook that is associated with this cluster.

Similarly, cluster 2 is also associated with a decrease in some of the variables that are decreasing in cluster 1 however to a lesser extent. The most notable difference between the two clusters is the high level of productivity, decreasing credit spreads and improving or appreciating currency. Cluster 3 and 4 are distinguished by their rising inflation, increasing GDP growth, rising credit spreads and decreasing yield curve levels. Next, we provide a more formalized definitions for these clusters.

5.2   REGIME CLASSIFICATION

In order to formalize the regime classification, the same regime definitions used by Dawsey (2014) and Emsbo-Mattingly et al (2014) is followed. These authors defined four economic regimes in their respective studies as recessionary, early cycle, mid-cycle and late cycle. Further discussion is done on the characteristics of each of these regimes and how they are translations of the different clusters observed in our study.

5.2.1   *Recessionary Phase*

In academia, a recessionary period is defined as a decline in GDP for more than two consecutive months (Emsbo-Mattingly et al, 2014). This period is associated with contraction in economic activities which precipitates declines in sales figures and corporate profits (Emsbo-Mattingly et al, 2014). From Table 5, this behavior in the economic variables is most evident from cluster 1. Furthermore, this cluster is associated with falling retail sales, GDP deflator and manufacturing order and sales.

Credit spreads were also observed to be wider on average during this cluster. The trade weighted index marginally deteriorates together with dividend yields. These observations were consistent with what would be expected under a recessionary regime. Thus we classify cluster 1 as a recessionary regime.

5.2.2 *Early Cycle*

Cluster 2 from Table 5 can be viewed as a cluster associated with improving conditions relative to cluster 1. In this cluster, we have consumer confidence and consumer spending improving as seen by an improved inflation and GDP averages, business confidence improves and production is rising. The most notable feature of this cluster is the steepening of the yield curve as seen by the increase in yield differential between the long end and the short end and narrowing credit spreads. In this case, we term cluster 2 as an early cycle in our regime classification. This regime will be characterized by improving economic conditions.

5.2.3 *Mid Cycle*

The third cluster from Table 5 distinguishes itself from cluster 1 and 2 by the positive outlook that it portrays within the economy. That is, in this cluster, inflation is rising, GDP is rising, productivity begins to slow down and yield curve levels begin to decrease. Furthermore, this cluster was the largest of the four clusters. In this case, we define this cluster as a mid-cycle cluster. This is consistent with the work of Emsbo-Mattingly et al (2014) where the authors defined the longest cycle in their model as mid-cycle.

5.2.4 *Late Cycle*

The Last cluster is associated with an overheating economy characterized by muted growths, spiking inflation numbers and constrained credit conditions. Furthermore, productivity and trade weighted index displayed significant deterioration while the increase in inflation averaged more than the other clusters. Growth as measured by the GDP was not indicative of any unique behavior. This cluster showed similar characteristics as those of a late cycle as given by Emsbo-Mattingly et al (2014). Next we consider how this framework can be used to forecast forthcoming regimes.

5.3 FORECASTING FORTHCOMING REGIMES

Given the regimes as defined in section 5.2, Table 6 on the next page provides a formalized classification of these regimes. This section illustrates how this result can be used for forecasting purposes.

In Table 6, we have the different regimes that have been translated from the clusters. Furthermore, the last two columns provides an illustration on how the model can be used for forecasting of upcoming regimes. This model follows from the work of Dawsey (2014).

Table 6: Regime Classification and definition

	Cluster Centers (Standardized Values)				Current State	
	Recessionary	Early Cycle	Mid-Cycle	Late Cycle	Current Period	Closest Match
CPI	−0.39	−0.10	0.06	0.53	0.20	*Late Cycle*
GDP	−0.36	−0.05	0.22	0.00	0.40	*Mid − Cycle*
Productivity	0.04	1.82	−0.47	−0.47	0.10	*Recessionary*
Changes in GDP Deflator	−0.10	0.16	−0.19	0.36	0.08	*Recessionary*
Manf. Order and Sales	−0.02	0.08	−0.05	0.04	−0.20	*Mid − Cycle*
Retail Sales	−0.28	−0.11	0.14	0.12	−0.15	*Early Cycle*
Current Acc. to GDP	0.14	0.12	−0.03	−0.16	0.20	*Recessionary*
Real Trade Weighted Index	−0.04	0.06	0.05	−0.15	−0.30	*Late Cycle*
Credit Spreads	0.89	−0.16	0.23	0.20	0.50	*Mid − Cycle*
PE Ratio	0.27	0.14	−0.24	0.09	0.4	*Recessionary*
Dividend Yields	−0.27	−0.08	0.16	0.00	−0.10	*Recessionary*
Changes on M1 money supply	−0.13	−0.17	−0.08	0.42	0.05	*Mid − Cycle*
Changes in Bond Spreads	0.66	−0.36	−0.13	−0.07	−0.50	*Early Cycle*
Current Period distance from cluster centers	1.62	2.01	1.20	1.25		

Column 5 in Table 6 provides an illustration of how a new data point, T_n where T_n is a vector that contains data measurements for the different economic variable and can be classified into one of the regimes. That is, assuming that time point T_n in column 5 is a future forecasted period which we require to know the regime it characterizes. This is done by calculating the distance between the regime cluster centers and the scaled data points for time period T_n. The distance between time period T_n and the regime cluster centers is given by the last row of Table 6. In order to obtain the values in the last row of Table 6, we calculated the Euclidian distance between the data points in column 5 and the regime centers. that is,

$$Distance\ from\ cluster\ center = \sqrt{\sum_{j}\sum_{i=1}^{n}(x_i - \overline{x}_{ji})^2} \tag{38}$$

where x_i is the centered macro-variable forecast data point and $\overline{x}_{ji}$ is the cluster centre for variable i in regime j,

$j \epsilon \{Recessionary,\ Early\ Cycle,\ Mid-Cycle,\ Late\ Cycle\}$.

From this calculation, it is clear that the time period T_n can be characterized as being in a Mid-Cycle regime since it has the smallest distance to cluster centers.

The fuzziness of any time period can be seen by the features of the different regimes by considering each regime variable in isolation. That is, we can look at each variable value in period T_n and decide which regime cluster center is the closest match. This is given by the last column of Table 6. It is evident that time period T_n exhibits regime qualities of all the different regimes to some degree. This result is more realistic than that which would result if the time points were forced to fully belong to a unique regime with full membership

The explanation given in the preceding paragraph serves as an illustration on how this model can be used to forecast expected regimes given forecasts of the macroeconomic variables. That is, the forecasting of upcoming regimes follows two stages. Firstly, a forecast of the economic variables is required. These forecasts are then used in the clustering process as shown through the process in which period T_n was classified into a regime.

This is a very attractive framework since many studies have shown how easy it is to forecast economic data and firm earnings relative to forecast of asset returns (Pesaran, Schuermann and Smith, 2008). Thus, for forecasting purposes, one would first establish a forecast of the economic variable for any required period . For each forecasted time period, the distance between that time period and cluster centers would be calculated to establish the expected regime.

This however is one way we can go about forecasting forthcoming regimes. Alternatively, we can use the principal components in a similar manner we used the original variables. That is, the cluster analysis produced cluster centers from principal components. When

Figure 9: Regime classification over the period under consideration

we have forecasts of the economic variables, we can simply perform a PCA decomposition and use the resultant principal components to calculate the distances to PCA cluster means like we did in Table 6 where the first column would be principal components.

We conclude this section by providing the Regime-Classification model for the study period under consideration. This is given by Figure 9 above.

From Figure 9 above, we have constructed the sequential progression through the different regimes over the study period. The next section, considers the behavior of the different asset classes across the regimes.

5.4 RISK AND RETURN ACROSS ECONOMIC ENVIRONMENTS

We initiate our empirical analysis with the risk and return behaviour of the asset classes in the different regimes. Table 7 at the top of page 55 provides an abstract of correlations between several asset classes over the full sample and in the different regimes.

On a full sample bases, we found that the average correlation of equity vs bonds, equity vs cash, equity vs gold and equity vs commodities were 0.34, −0.48, 0.51 and 0.23 respectively. We note however that the correlation between cash and equity becomes positive dur-

Table 7: Key correlations across different economic regimes. Sample period 1993-2014

	Equity − Bonds	Equity − Cash	Equity − Gold	Equity − Com
Panel A				
Full Sample	0.34	−0.48	0.51	0.23
Panel B				
RECESSION	0.67	0.24	−0.11	0.12
EARLY CYCLE	0.04	0.04	0.13	0.17
MID-CYCLE	0.20	−0.21	0.11	0.39
LATE-CYCLE	−0.10	−0.19	0.17	0.26

ing early cycles which is contrary to the negative average correlation observed over the full period. The correlation between equity and bonds is positive over the full sample while during late cycles, this correlation becomes negative. Furthermore, the correlation between gold and equity become significantly negative during recessionary periods while it stays positive in the other cycles.

These findings are similar to empirical findings found by other studies on the JSE market. In a masters theses titled "Evaluation of Gold as an Investment Asset", Pule (2013) demonstrated how gold reduces systematic risk when added to the portfolio. Even though the focus of the author was not on regime analysis nor a comparison of asset class performance, his results demonstrated diversification potential of the Gold asset class. In a similar manner, Bodington (2014) illustrated in his thesis titled "Gold in the South African market: A safe haven or hedge" the ability of gold to act as a hedging asset class during bad financial times. Clearly, from Table 7, we find that gold may potentially provide some diversification potential during recessionary regimes.

Furthermore, from Table 7, the correlation between equity and bonds becomes significantly negative during late cycles while significantly positive during recessionary periods. This observation is also consistent with the vast body of empirical results which illustrates how bonds lose their diversification ability during bad times (Kollar (2003), Chin (2013) etc.). The findings on the correlation behaviour of the asset classes relative to the equity asset class are simimilar to those found by van Vliet and Blits (2009).

Table 8 at the top of page 56 provides the annualised volatilities for the different asset classes across the different economic regimes and for the full sample. On average, we have found that risk tends to be heightened during the recessionary regime relative to the other regimes. The risk of equities in particular tends to be higher during the recessionary regime while it remains stable during the early and

Table 8: Risk of each asset class as measured by the annualised volatility. Sample period 1993-2014

	1993-2014	RECESSION	EARLY CYCLE	MID-CYCLE	LATE CYCLE
Cash	1.04%	0.52%	0.13%	0.13%	0.10%
Bonds	8.69%	8.67%	9.93%	7.08%	8.21%
Equity	19.19%	23.13%	20.85%	15.16%	11.2%
ILB	4.48%	2.27%	2.27%	3.76%	4.25%
Property	16.21%	16.85%	9.93%	13.11%	19.29%
F. Cash	16.21%	15.57%	12.84%	15.64%	14.27%
F. Bonds	15.83%	15.52%	11.91%	14.62%	15.39%
F. Equity	16.91%	18.82%	14.68%	13.56%	11.99%
Gold	20.11%	23.30%	12.51%	17.26%	23.54%
Commodities	17.41%	16.91%	13.81%	13.99%	16.69%

mid-cycle. It is also interesting to note that under the full sample, the equity volatility of 19.19% is less than the 23.12% volatility observed during recessionary period while it is higher than the 11.2% observed during late cycle. The implication of this result is that if we do not account for time-varying risk embedded within the different regimes, risk may be understated during bad times but over stated during good times.

The time-varying risk profile in this case would be propagated by the increase in risk in asset classes during the bad times. In particular, the variation in risk across asset classes and the increasing correlation between equity and other asset classes previously mentioned further displays evidence of the time-varying risk nature of asset classes. It is also worth highlighting how the volatility of commodities and gold increases during the recessionary period and late cycle while it is fairly stable during early and mid cycle. Bonds and inflation linked bonds showed limited time-variation across the different regimes.

Table 9 on the next page provides the average returns for each of the asset classes through the different regimes. It is observed that the returns of equity are highest during recessionary and late cycle periods. Whilst initially it may seem unintuitive that the returns are highest during the recessionary regime, we highlight later that equity markets usually crash in anticipation of a recession and recover from this low point during a recessionary period, thus resulting in somewhat higher returns during the economic recession. On the other hand, bonds perform relatively better during early and mid-cycle. This finding on these asset classes is also consistent with the works of Van Vleit and Blitz (2009), Ang and Bekaert (2002), Sa-Aud, Shilling and Tiwari (2005). Property on the other hand out performs all asset

Table 9: Annualised returns of asset classes for each regime. Sample period 1993-2014

	FULL SAMPLE	RECESSION	EARLY CYCLE	MID-CYCLE	LATE CYCLE
Cash	9.86%	10.92%	12.71%	10.83	7.86%%
Bonds	12.67%	8.15%	17.67%	17.18	9.27%
Equity	15.70%	28.36%	25.15%	10.32%	15.82%
ILB	11.93%	6.49%	13.86%	11.34%	11.67%
Property	20.97%	21.46%	11.51%	19.90%	26.80%
F. Cash	8.82%	−9.39%	25.15%	18.92%	9.54%
F. Bonds	11.39%	−7.30%	25.34%	28.50%	8.38%
F. Equity	13.77%	12.20%	28.72%	10.18%	15.85%
Gold	12.71%	0.297%	28.53%	20.05%	13.53%
Commodities	5.89%	−1.78%	25.34%	−1.07%	1.24%

classes during the late cycle while foreign bonds and gold perform the best during mid-cycle.

5.5 REGIME - BASED ASSET ALLOCATION

The underlying philosophy that underpins regime-based asset allocation is that, if expected economic conditions can be assumed to persist with a strong link to asset class performance, then conditioning asset allocation on these environments should be warranted. In its embryotic form, regime-based allocation instigates a dynamic process for strategic asset allocation. Naturally, this form of asset allocation is long-term based (Lyngby, 2014). That is, the choice of a portfolio composition is based on long-term views of asset class performances.

The aim is to provide portfolio managers with a competitive edge that allows them to exploit favourable economic conditions while reducing drawdown potential during adverse economic conditions. Thus, regime-based asset allocation is designed in a flexible manner which makes provision for time-varying risk factors. Merton (1973) illustrated in his work the vulnerability of optimal portfolios to uncertain future investment environments. Hence conditioning portfolio construction on expected investment environments or rather regimes, provides a more attractive framework to mitigate time-varying risk.

The analysis that follows in this section assumes an absolute investment environment. In the optimisation process, the assumption of long-only positions is made. We begin our analysis of regime-based asset allocation with an analysis of optimal portfolio composition for each regime separately. This will be followed by an in-sample performance analysis of the allocation strategy.

Table 10: Optimal Weights across regimes (%)

Asset Classes	Recessionary	Early Cycle	Mid-Cycle	Late Cycle
Cash	0.13	0.00	0.13	2.45
Bonds	0.13	2.30	12.18	0.15
Equity	53.13	12.50	1.23	19.00
Inflation Linked Bonds	0.00	9.40	2.05	0.00
Property	44.07	0.10	35.25	48.90
Foreign Cash	0.07	4.40	1.88	0.70
Foreign Bonds	0.07	10.20	1.00	0.00
Foreign Equity	1.93	4.90	3.33	2.55
Gold	0.00	21.20	33.35	7.25
Commodities	0.47	35.00	9.63	19.00

5.5.1   *Determining the optimal asset allocation in each of the economic regimes*

Under the proposed regime-based asset allocation method, prior knowledge of a preferred portfolio mix in each regime is a key requirement. In order to achieve this requirement, it necessitates the exploration of the optimal portfolio composition in each of the regimes separately.

As mentioned in the methodology section, the optimization procedure that was followed is that proposed by Almgren and Chris (2004). The use of this technique can be motivated by the fact that return magnitude estimates usually have estimation error. Managers at best are likely to better forecast the direction of returns and perhaps the rankings of returns across asset classes. One optimization technique that utilizes ordinal return data is the technique proposed by Almgren and Chriss (2004). Additionally, Almgren and Chriss (2004) use a centroid signature that transforms the linear rankings of returns into a more realistic shape that more closely approximates the return behavior at the tails of the return vector. Thus, the optimization is driven by ordinal return signals rather than the magnitude of point estimates. The view is taken that it is more important having knowledge of the ordinal structure of the asset classes than the magnitudes of returns.

Table 10 above provides the optimal weights obtained from the optimization under each regime.

A few interesting observations can be made from Table 10. Firstly, it is observable that no single asset has the highest weights across all regimes. This can be seen by the different weights assumed by the asset classes across the regimes. In this regard, no single asset class demonstrates dominance over all regimes. This observation is consistent with findings from research by a number of authors (Sa-

Aadu, Shilling and Tiwari, 2005; Eychenne and Martinetti, 2011; Kollar, 2013). In order to elaborate more on this result, we chart the ranked weights of the portfolio composition in each regime . This is given by Figure 10 on page 60.

In Figure 10, we observe that during the recessionary regime, the portfolio overweights local equity followed by property. This result provides some insightful results on the behavior of financial markets relative to the real economy. That is, the financial market is viewed as a leading economic indicator in most econometric models (Auret and Golding, 2012). The result observed in this thesis is supported by results from other studies since the implications of an overweight in equity during recessionary periods implies that the markets would have already priced in a recessionary period prior its occurrence.

Furthermore, several academic authors have shown how market down turns were followed by recessions (Islam and Verick, 2010; Kannanm, Scott and Terrones, 2012). This is also in line with accommodative policy implementation by many governments during recessionary period which are positive for equity markets (Kannanm et al, 2012). Within our framework, this allocation is consistent with expected performance of the asset classes within this regime. That is, during the recessionary regime that has been defined in this thesis, we had observed that local equity outperformed other asset classes.

The weighting changes during the early cycle and we observe that the portfolio overweights commodities followed by gold. Unlike the recessionary regime, the early cycle regime is less concentrated. In this case we observe a more diversified portfolio with investment in 8 out of the 10 asset classes. Very interestingly, we find that the weight on equity and property change significantly. The weighting on equity is reduced by more than 50% while property is reduced to almost zero. The allocation gravitate more towards currency driven asset classes.

During the mid-cycle, the weights gravitate back to property which obtains the highest weight followed by gold and local bonds. During this regime, we had observed that the best performing assets were foreign assets led by foreign equity. However due to the 25% restriction on foreign allocation, we had total foreign allocation accounting for 19.5% within this regime. This restriction allowed for an increased allocation to other assets which were not necessarily the best performing assets within this regime.

During the late cycle, property maintains its ordinal position as seen during mid-cycle with a relatively increased weighting. This is consistent with the fact that property was shown to be the best performing asset during this regime. In Figure 11 on page 61, we provide an analysis of marginal risk contribution of the different asset classes as measured by the Expected Tail Loss (ETL). This measure is generated using the proportional allocation of the different asset classes.
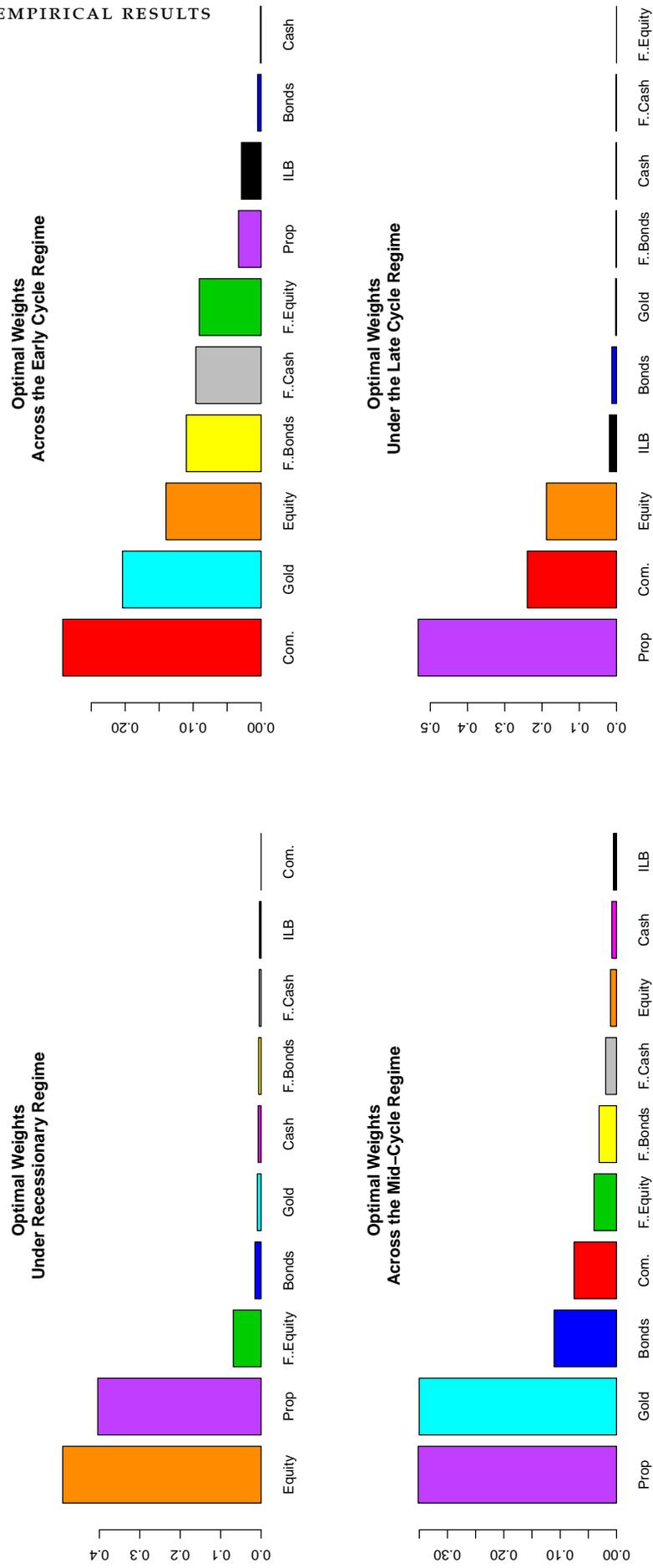
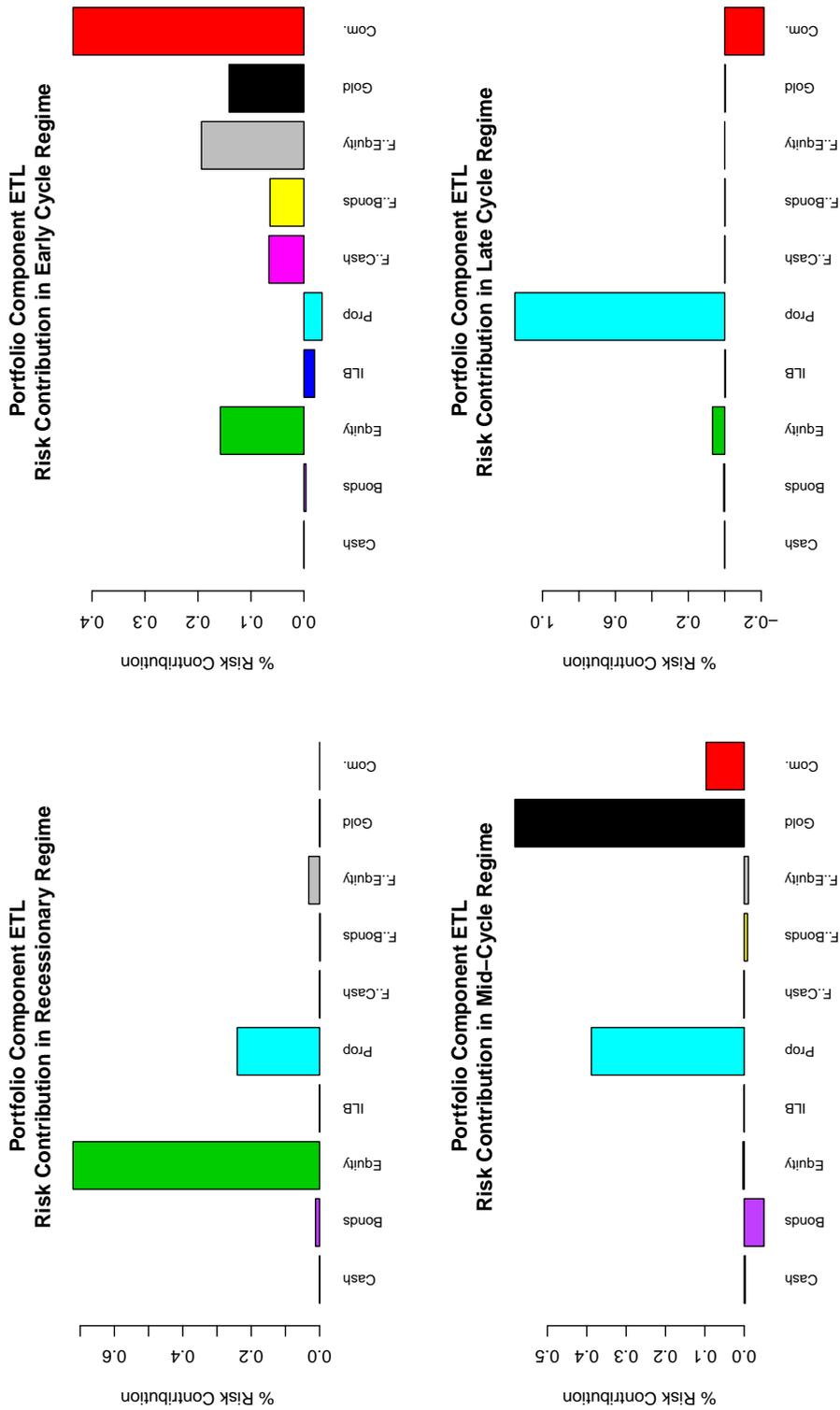Figure 10: Portfolio Composition across the different economic regimes

Figure 11: Marginal Risk contribution to the portfolio in each regime

This measures the additional risk to the overall portfolio that would be incurred if we were to increase our allocation to the assets classes or if a major negative shock occurs.

A positive value implies an increase in risk while a negative value implies diversification potential. For our purposes, we have used the Conditional Value at Risk as a measure of risk contribution to the portfolio from the different asset classes.

In the top pannel of Figure 11 on the previous page, we find that the highest contributors to portfolio risk as measured by the ETL are equity and property. This is consistent with the observed concentration in weights on these asset classes during this regime. Within the Recessionary regime, it was observed that all assets were either contributing positively to the portfolio risk or having no effect at all. This however changes in other regimes.

In the early cycle, the highest contributor to portfolio risk was gold. However, it was observed that local bonds, equity and property showed marginal diversification potential. That is, these assets have negative contribution to portfolio risk which is a characteristic of diversification in this context. Similarly during mid-cycle and late cycle, bonds exhibit a strong diversification potential since they have a negative marginal risk contribution. These results are well supportive of the empirical observations provides in section 5.3 where we considered the behavior of the different asset classes across the regimes. The implications of this result is that in the event that a portfolio manager seeks to reduce concentration within the portfolio at any point in time, the observed assets with negative risk contribution in each regimes are where allocation should gravitate to.

In the next section, we assess the performance of the proposed asset allocation strategy against an equally weighted portfolio and a buy-and-hold strategy. In the process of assessing the performance of the regime-based asset allocation strategy, rebalancing is done at the beginning of each regime. We further assume that we have 100% foresight of when regimes change hence rebalance accordingly. The choice of an equally weighted portfolio and a buy and hold as comparable portfolio's was motivated by other studies that compared regime-based strategies to these strategies.

5.6   IN-SAMPLE PERFORMANCE ANALYSIS

This section provides in-sample performance results for the proposed regime-based asset allocation strategy. As mentioned in the previous section, the in-sample performance results of the regime based strategy will be compared to in-sample performance results of an equally weighted portfolio and an optimized buy-and-hold strategy.

The main objective of this section is to illustrate the merits of a regime-based strategy through an in-sample performance analysis
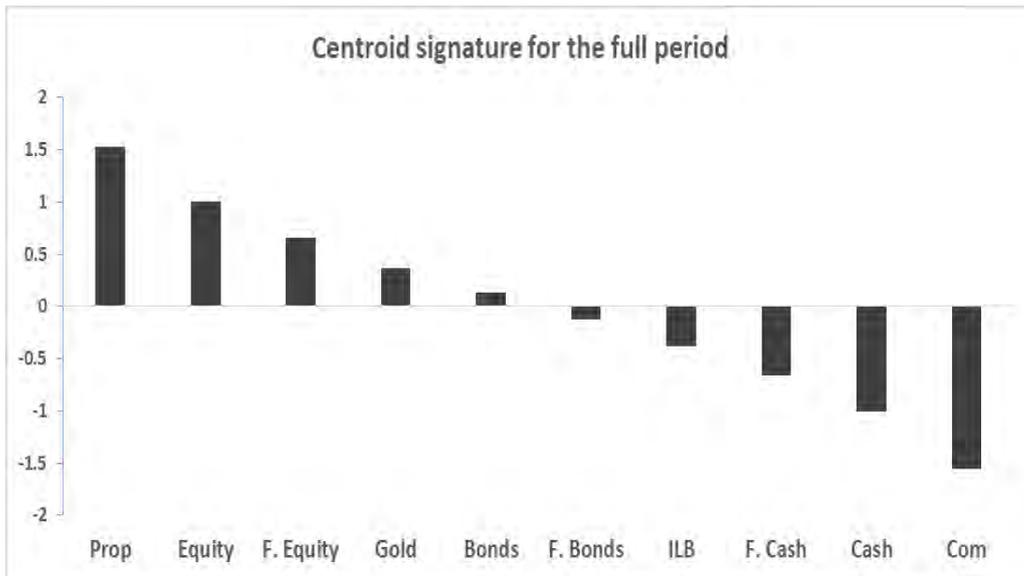
Figure 12: Centroid signature used in the optimisation for the buy-and-hold strategy

relative to alternative allocation methods. That is, one could consider an equally-weighted portfolio as an allocation strategy or rather invoke a buy-and-hold strategy without taking a view on the state of the economy. In this case, if these strategies, on an in-sample basis, outperform the proposed regime-based allocation method; it would yield less of a motivation for an out-of-sample performance test.

Under the regime-based strategy, the portfolio is rebalanced to the respective weights of the regimes shown in Table 10 on page 58. The assumption that the regime switch or regime change occurs at the rebalancing point is made in this case. This is a very strict assumption since it implies that the portfolio manager ought to know when economic environments change from one regime to the other. However, for the sake of illustration, this assumption provides no hindrance to the analysis.

In our analysis, we further assume that managers know when the regime change occurs. This implies that managers know precisely when the portfolio composition should change and how to allocate the assets as per the strategy. This assumption is however unrealistic in practice but in order to facilitate a justified comparison to previous studies, this setting is more appropriate.

In order to make the buy-and-hold portfolio strategy comparable to the regime-based portfolio strategy, the centroid optimal portfolio will be used. That is, we use the same centroid optimization methodology that was implemented for each regime for the buy-and-hold strategy. This helps standardize the optimization technique used for the different strategies which in turn enables us to compare their performances.

Table 11: Optimal weights for the buy-and-hold strategy

Asset Classes	Weights
Cash	0.00
Bonds	0.00
Equity	0.19
ILB	0.08
Property	0.45
F.Cash	0.00
F.Bonds	0.12
F.Equity	0.09
Gold	0.07
Com	0.00

Figure 12 at the top of the privious page provides the centroid signature that was used in the optimization for the buy-and-hold strategy.

The centroid signature provided in Figure 12 essentially transforms the linear nature of the return rankings into a signature that is more similar to how returns typically portray more extreme behavior in the tails. This was based on the ordinal ranking of the average returns of the different asset classes over the full period. This process is consistent with the observed cumulative returns of the asset classes provided in section 3.2.2. The resultant weights from the optimization can be seen from Table 11 above.

Figure 13 on the next page provides cumulative returns for the 3 strategies under consideration. The black time series represents cumulative returns from the buy-and-hold strategy. The red time series on the other hand provides the cumulative returns from the equally weighted portfolio while the green time series provides the cumulative returns for the regime-based strategy.

It is clear from Figure 13 that the regime based strategy outperforms the buy-and-hold strategy and the equally weighted portfolio on cumulative basis over the study period. This result is again consistent with the body of literature that has shown the superiority of a regime-based strategy over an equally weighted portfolio or a buy-and-hold strategy (Bulla et al, 2010). However we note that this hardly surprising as the analysis was performed in-sample, but does highlight the potential that can be gained by forecasting forthcoming regimes accurately. Table 12 on page 66, provides the performance summary from the different strategies.
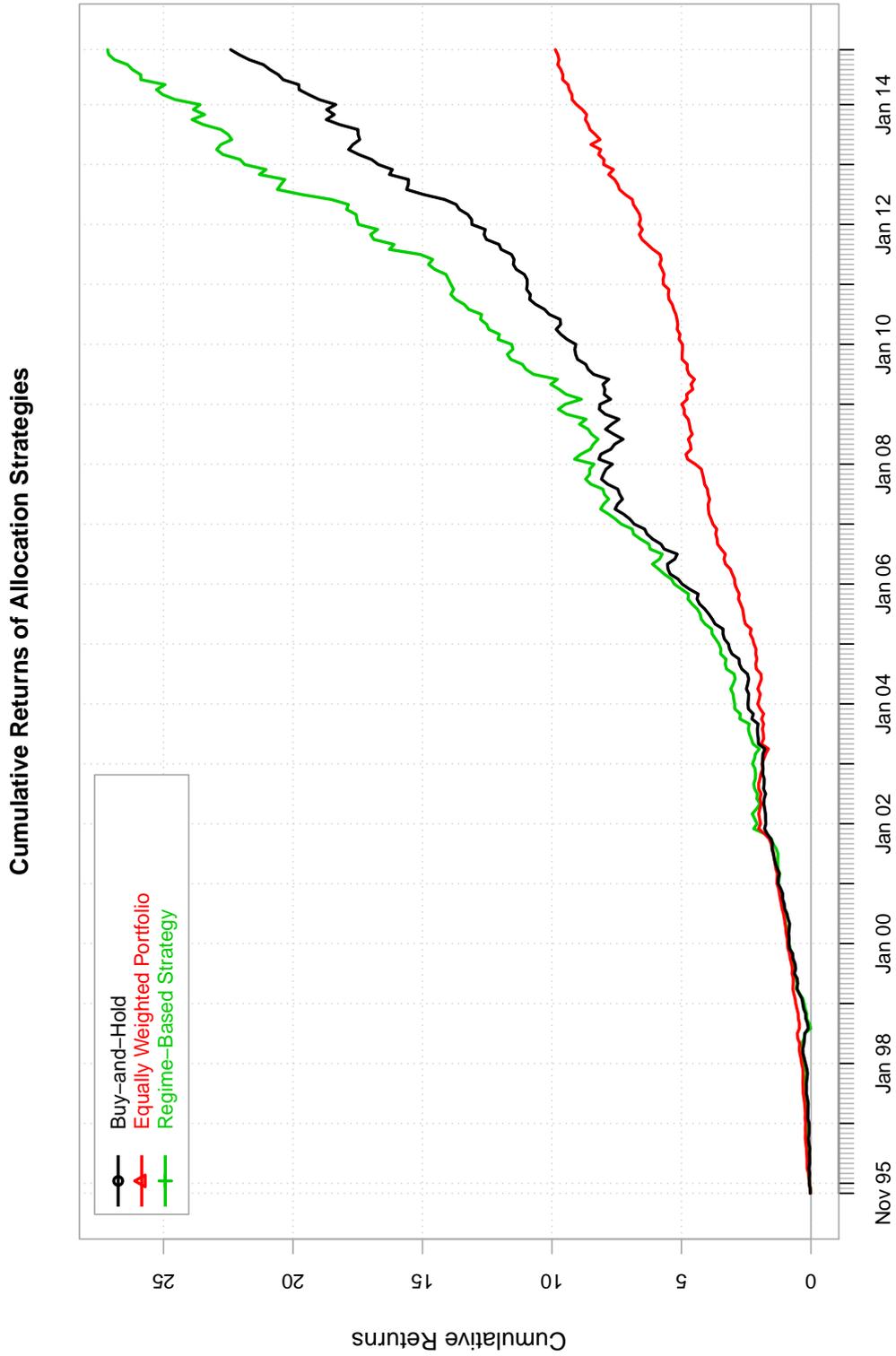
Cumulative Returns of Allocation Strategies

Figure 13: Performance summary of the different strategies

Table 12: Summary statistics for the strategies

Strategy	Ann. Return	Ann. Std	Ann. SR
Regime-Based	19.27	10.35	1.86
Buy-and-Hold	17.34	8.02	2.16
Equally Weight	13.25	7.22	1.84

Table 12 above provides a summary of the in-sample return performances for the 3 strategies. The first column of the table provides the strategies under consideration. The second column provides the annualized returns while the third column provides the annualized standard deviations. The last column provides the annualized Sharpe Ratio for each of the strategies.

Table 12 quantifies the results portrayed in Figure 13 of this section. We observe that the regime-based strategy has a higher annualized return (19.27) relative to the buy-and-hold (17.34) and an equally weight (13.25) portfolio. However, when we consider the Sharpe Ratio, we observe that the buy-and-hold strategy provides better risk adjusted returns relative to the regime-based strategy and the equally weighted strategy. This comes as a result of a higher standard deviation of the regime-based strategy which dampens the risk adjusted returns.

Bourachnikova and Yusupov (2011) argued against the use of variance as a measure of risk with respect to portfolio management. Following the argument by Bourachnikova and Yusupov (2011), the information carried in the Sharpe Ratio becomes distorted. These authors proposed the use of semi variances as a better proxy while Sortino and Prince (1994) had initially proposed the use of downside deviations resulting in what is known as the Sortino Ratio. In these cases, the point being made is that investors care most about negative deviations rather than the overall average deviations. We consider the use of the Sortino Ratio rather than the Semi-Variance method.

The concept behind the Sortino Ratio is similar to that of the Sharpe Ratio. That is, both the Sharpe Ratio and the Sortino Ratio measures the risk adjusted returns of a portfolio or individual asset. The distinguishing factor between these two approaches is that while the Sharpe Ratio treats upside and downside deviations equally in its use of the variance, the Sortino Ratio considers deviations given by returns falling below a defined minimum acceptable return. That is, instead of using the variance of the returns, we use downside deviations in the calculation of risk adjusted returns.

Sortino further proposed an upside potential ratio which is computed in a similar manner as the Sortino Ratio but uses upside returns in the numerator rather than average returns. The upside potential ra-

Table 13: Summary of Risk Performance Ratios

Strategy	Sortino Ratio	Upside Potential	Max Drawdown
Regime-Based	1.484	2.035	0.155
Buy-and-Hold	1.443	1.802	0.115
Equally Weight	1.119	1.515	0.126

tio, Sortino Ratio and Drawdown potential provides strong pillars to assess performance in a given strategy.

Table 13 above provides the measure for the Sortino Ratio, Upside Potential and Maximum Drawdown's which compares the 3 strategies.

Table 13 provides a different outcome than that observed when using standard deviation as a measure of risk. That is, when using the Sharpe Ratio analysis, the buy-and-hold strategy outperformed the other strategies. However, if we consider the use of downside deviations as a measure of risk, we find that the regime-based strategy provides better performance compared to the other strategies.

In this case, we observe that the regime-based strategy compensates for almost 1.5 times for downside risk incurred in the portfolio. Furthermore, the regime-based strategy has the highest upside potential relative to the other strategies where the upside potential is given by deviving upside deviations by the downside deviations.

These results demonstrate the potential outperformance of the regime-based strategy. It further demonstrates the adequacy of the proposed regime classification method in capturing the distinctive regimes within the study period.

It is also worth looking at the rolling performance analysis of the strategies. This is provided in Figure 14 on the next page.

Figure 14 provides $24 - month$ rolling returns for the different strategies. Evidently the regime-based strategy outperforms the two comparable strategies on rolling bases. The only period that the strategy provides a lack of performance relative to the buy-and-hold strategy was $2005 - 2006$ period and the $1997 - 1998$ period. Post 2007, the regime-based strategy exhibits strict outperformance relative to the other strategies. Figure 15 on page 69, provides the rolling Sortino Ratio's for the different strategies. This further shows how the regime-based strategy provides better compensation for drawdown risk associated with the strategy. These results illustrate the benefits that could be achieved through the consideration of regimes when making asset allocation decisions. Furthermore this warrants the regime classification method followed in this thesis with a potential for further investigating the robustness of the method.

On an ending note to this analysis, we provide a monthly relative performance measure of our strategy relative to the mentioned
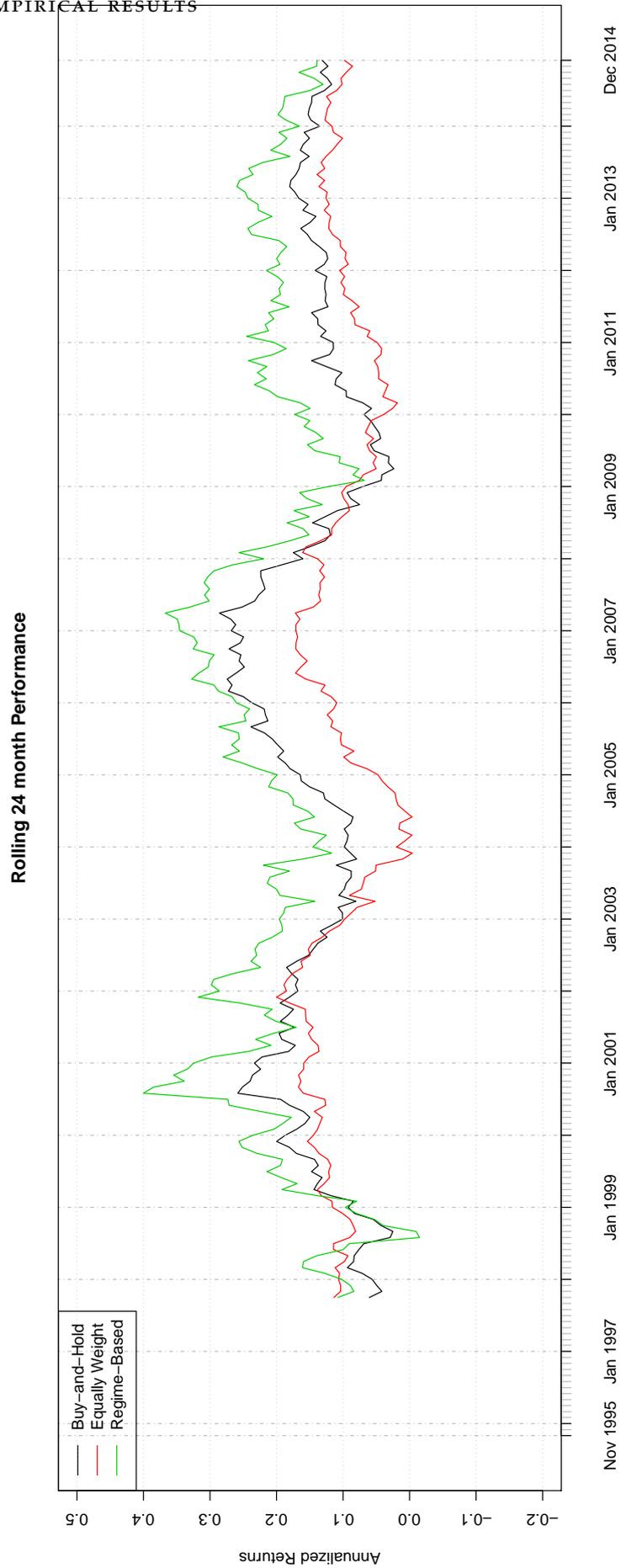
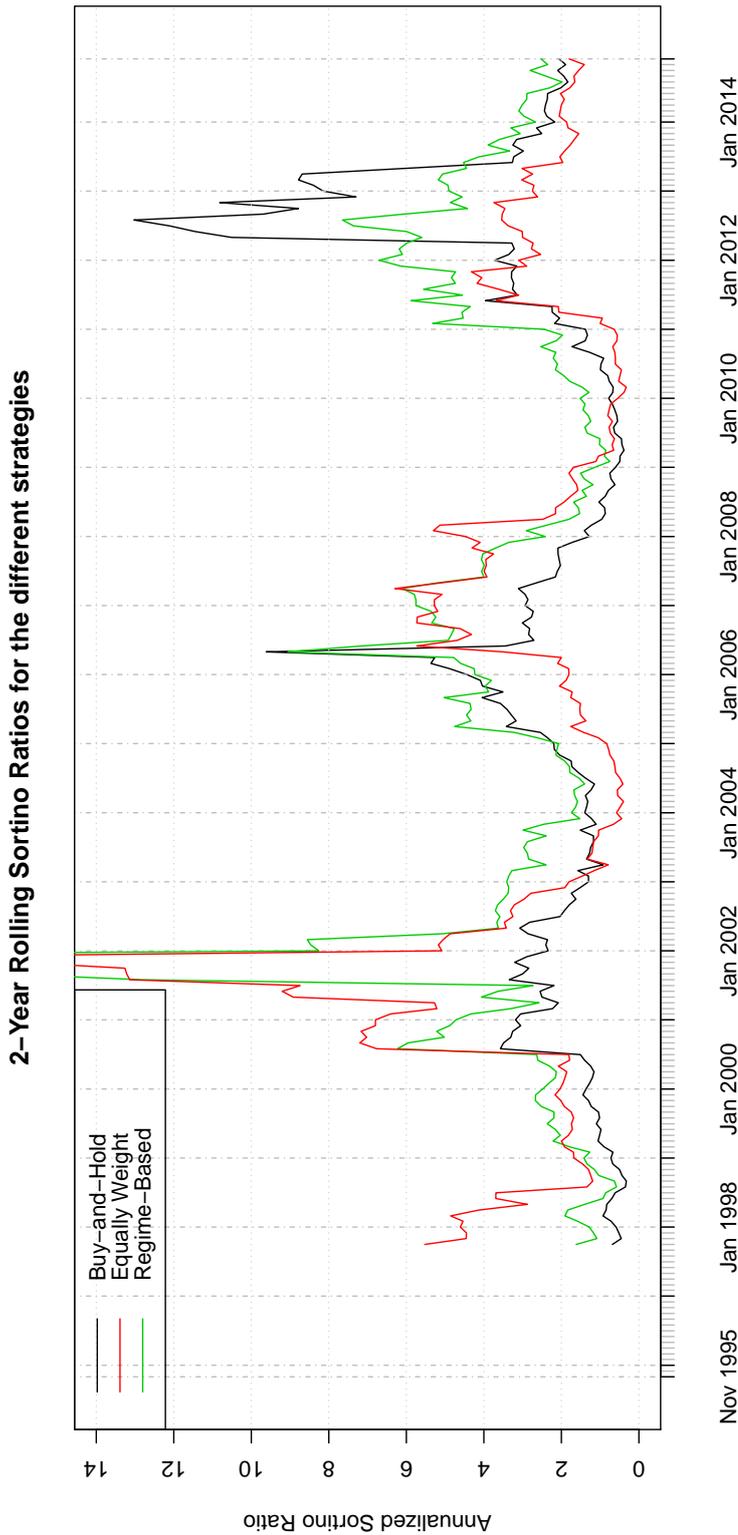Figure 14: Two-year Rolling period performance analysis

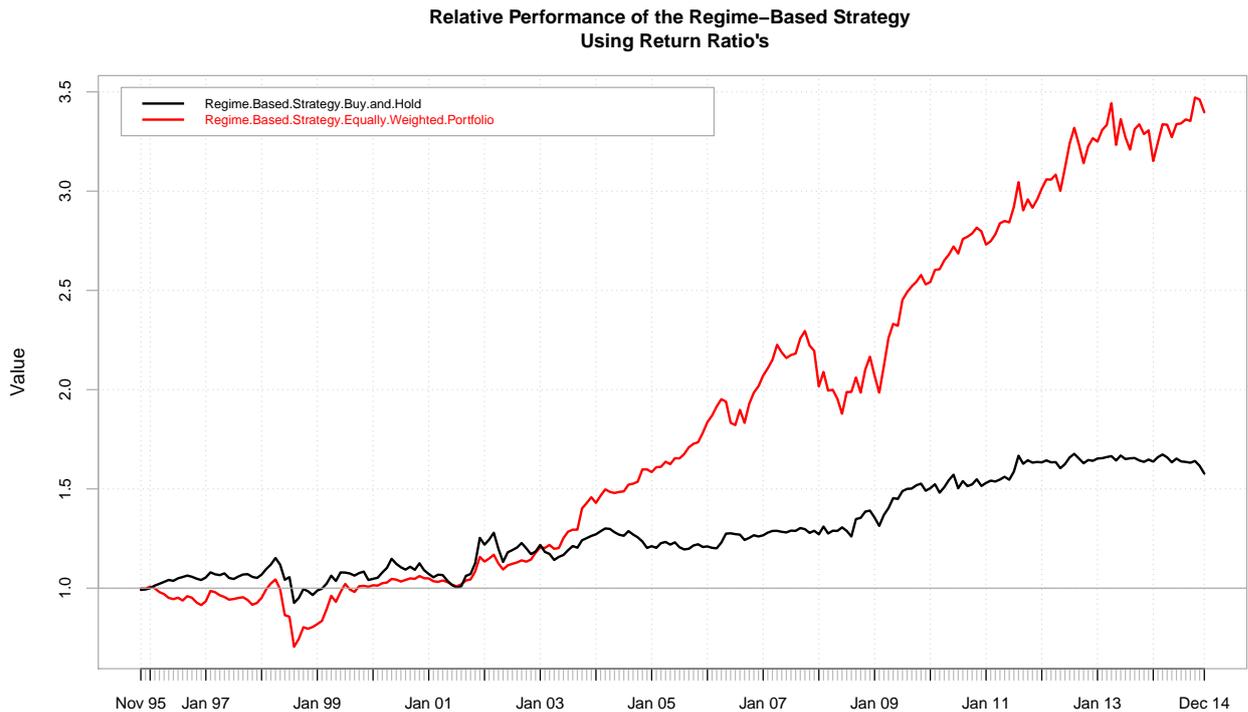Figure 15: 2-Year Rolling Sortino Ratios for the respective strategies

Figure 16: Relative performance of the regime-based strategy

comparables. This is done by dividing the monthly returns generated from the Regime-Based portfolio by those generated by the other two respective strategies. This is given by Figure 16 above.

In Figure 16, the black time series providing the ratio of returns between the regime-based portfolio and the buy-and-hold. The red time series on the other hand provides the ratio between the regime-based portfolio and the equally weighted portfolio. From this figure, one would expect the ratio to be equal to one in cases where the buy-and-hold and equally weighted portfolios generate returns that are equal to those of the regime-based portfolio. In cases where the regime-based portfolio achieve returns greater than the other two portfolios, the ratio will be greater then one. If the regime-based portfolio is outperformed, than the ratio will be less than one.

Figure 16 provides a simplified summary of the relative performance of the regime-based strategy to the equally-weight and buy-and-hold strategy. The regime-based strategy demonstrates outperformance on both the comparable strategies. The period 1998–1999 is the only significant outlier period where the strategy underperforms. This may be associated with the period where the drawdowns from the regime-based strategy exceeded those of the other two strategies. However post this period, the regime-based strategy provides superior performance over the comparable strategies.

The results obtained in this section provide some comfort on the in-sample performance measure of the regime-based strategy. More importantly, we have established that the regime-based strategy proposed in this thesis significantly outperforms an equally weighted strategy. This stems from the fact that its outperformance over an equally weighted strategy places value on the asset allocation process that is initiated by imposing views an expected regime. Thus, rather than simply equally weighting a portfolio or using an optimised buy-and-hold strategy, the effort of constructing regimes warrants outperformance. Secondly, even though the strategy did not completely outperform the buy-and-hold portfolio on all the performance measures considered in this section, we found comfort in knowing that it performed just as good in cases where it underperformed it's historic levels implying a very attractive allocation procedure.

RELATIVE PERFORMANCE ANALYSIS

This chapter provides a further analysis of the performance statistics observed from the regime-based strategy. However, it may not be easy to interpret these statistics in isolation. Therefore, in order to give context to the interpretation of the empirical results of the regime-based strategy, it is worth contrasting it to results that were achieved by the largest pension funds over the same period. For this purpose, the Alexander Forbes Large Manager Watch (AFLMW) index will be considered. The AFLMW index is a survey that showcases the average performance of the largest 10 South African pension funds. This index dates back to 1999 and provides a relatively good index to contrast the performance that could potentially be achieved through accurate regime analysis based on our proposed model. Thus, the comparison will be done over the period 1999 to 2014 rather than the period 1993 to 2014. This implies that the calculated results of this chapter (such as returns etc) will differ from those observed in chapter 5. Admittedly so, the comparison may not be a fair one due to the foresight bias of the regime strategy. However, this analysis will provide some insight on the potential that could be achieved by the AFLMW index. That is, any differential in performance will be indicative of the potential that could be achieved and the added benefit that would result through attempts of incorporating regime changes in allocation strategies.

In order to achieve the above-mentioned goal, this section is segmented into two parts. Firstly, the analysis is focused on return performance summary and the risk measures given by the first four moments of a distribution, namely the mean, variance, skewness and kurtosis. The second section will focus on the risk metrics that most practitioners consider for performance evaluation. In particular, the second section will consider risk metrics such as downside deviations, drawdowns, Value at Risk (VaR) and Expected Shortfall (ES).

6.1 PERFORMANCE SUMMARY STATISTICS

In this section, we take a closer look at the performance statistics and first four moments that characterizes the return distribution generated by the regime-based strategy and the AFLMW return index for the period 1999 to 2014. These are given by Table 14 on the next page. The reader is reminded once again that these calculations are calculated over the period 1999 to 2014 hence the difference from those observed in chapter 5.
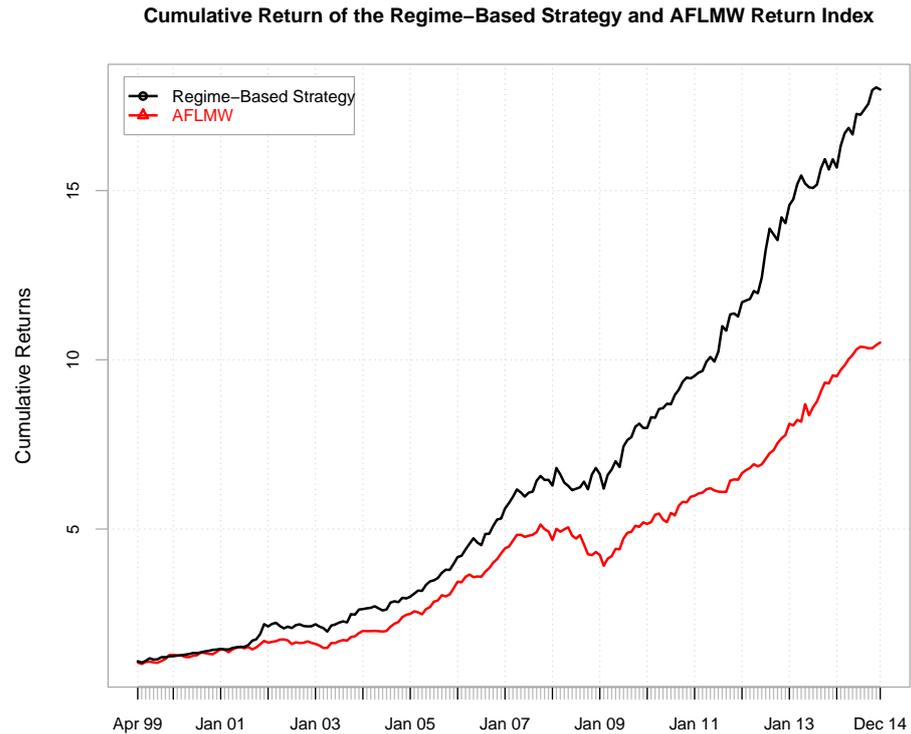
Cumulative Return of the Regime–Based Strategy and AFLMW Return Index

Figure 17: Cumulative Returns from the regime-based strategy and AFLMW

From Table 14 on the next page, it is observable that on an annualized basis, the regime-based strategy has yielded better returns (20.35) than what the AFLMW index achieved (16.11%). More specifically, if we assume that managers could have forecast the regimes using the proposed model, their average return would have improved by 4.24% per annum. This can also be seen through observing the cumulative returns from the two series. In Figure 17 above, the cumulative returns from the regime-based strategy and the AFLMW index are provided.

Table 14: Perfomance summary statistics for the period 1999 to 2014

Measurement	Regime-Based Strategy	AFLMW Index
Annualised Return	20.35%	16.11%
Annualised StdDev	10.78%	10.66%
Annualised Sharpe Ratio	1.89	1.51
Skewness	0.77	-0.03
Kurtosis	1.76	0.18

**Five Number Summary for the
Regime–Based Strategy and AFLMW Return Index**

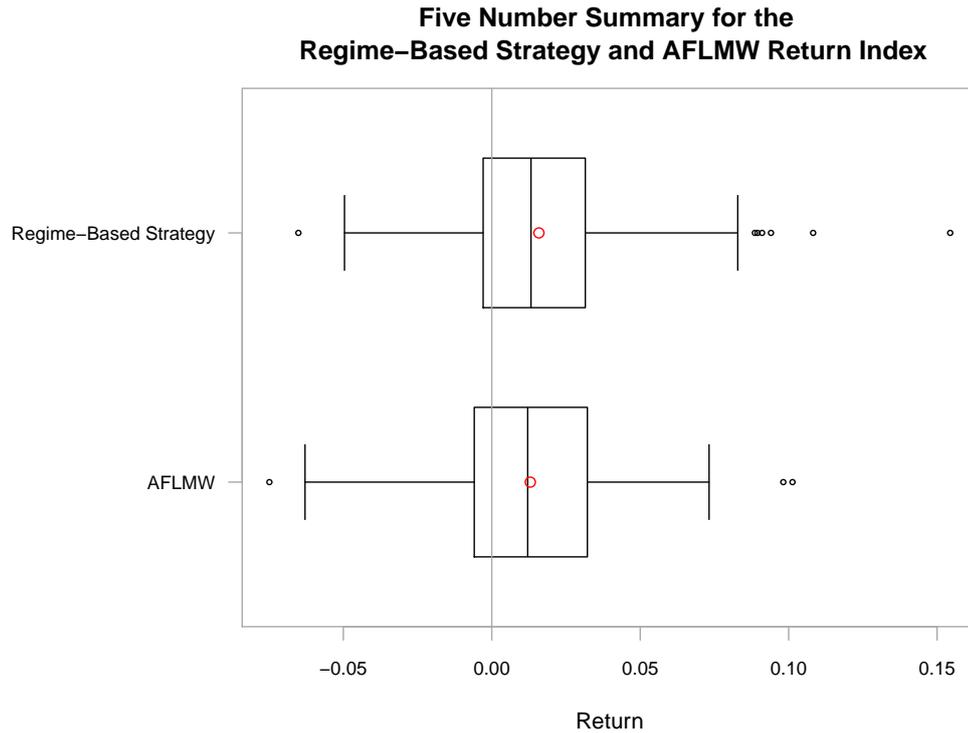Figure 18: Five number summary of the return series

From Figure 17, the black time series represents the cumulative returns from the regime-based strategy while the red time series represents the AFLMW cumulative return series. It is highlighted again that even though this may not be a fair comparison, it does provide information on the potential that could be achieved through the efforts of basing asset allocation decisions on regimes through the proposed method. Thus, from Figure 17, it can be argued that portfolio managers could obtain some benefits through the use of the proposed regime model and accurate efforts of forecasting.

In order to obtain a better understanding of the behavior of the returns around their long-run mean, a five number summary using the box-whisker plot is given in Figure 18 above.

Figure 18 above provides a very interesting view when comparing the distributional properties between the regime-based strategy and the AFLMW return series. Firstly, similar to the results from Table 14, it can be observed that the average return from the regime-based strategy distribution appears to be higher than that observed from the AFLMW return index. Secondly, when considering the dispersion around the mean and the interquartile range (spread between the upper and lower quartile), the regime-based strategy has a narrower "spread" than the one observed under the AFLMW index. The implication of this observation is that the returns under the regime-based strategy are relatively more predictable than those currently

observed from the AFLMW index. It can also be observed that most of the outliers in the regime-based strategy lie on the positive side of the mean. This is an attractive feature displayed by the regime-based method as it provides a framework that has a relatively predictable return structure while exposing managers to upside potential..

The annualized standard deviation from the regime-based strategy appears to be higher than that of the AFLMW index. From Table 14, the regime based strategy has an annualized standard deviation of 10.78% while the AFLMW index recorded an annualized standard deviation of 10.66%. On face value, one could argue that from these numbers, the regime-based strategy is more risky than the average allocation strategy employed by the average manager included in the AFLMW list.

However, the use of variance or standard deviation as a measure of risk has been argued as not being the best measurement of risk (Satchel and Sortino, 2001). One of the arguments that is often sited in this regard is that variance or standard deviation considers an average of upside and downside deviations in the return distribution (Satchel and Sortino, 2001). In practice however, many would argue that portfolio manager are primarily concerned with downside deviations. That is, it is more important to portfolio managers to know how much they could lose relative to how much they could gain. Thus, variance or standard deviation provide little information on the matter. The most commonly used risk metric in substitution of variance are drawdown measurements. Drawdown calculations measure pick-to-though declines within a given period of analysis (Satchel and Sortino, 2001). Figure 19 at the top of the next page provides a comparison of monthly drawdown's between the regime-based strategy and the AFLMW return index.

Clearly from Figure 19, downside deviation of the AFLMW return series larger more than those observed from the regime-based strategy. Even more interestingly, it can be observed from Figure 19 that during regime shift periods and adverse economic conditions, the regime-based strategy provided more downside protection than what managers appear to have had. That is, from Figure 19 above, during the Asian crisis of 1998-2000 period, the AFLMW index lost close to 5% on average, while the regime-based strategy had close to zero drawdowns. Furthermore, during the global credit crises of 2008-2009 period, the AFLMW index shows that managers had drawdowns of more than 20% while the regime-based strategy lost less than 10%. Thus, even though the standard deviation of the regime-based strategy may have appear to be higher than that of the AFLMW index, the deviation are arguably more on the upside than on the downside. This is well supported by the observed outliers of the regime-based strategy that were on the positive side of the mean in the box-and-whisker plot of Figure 18.
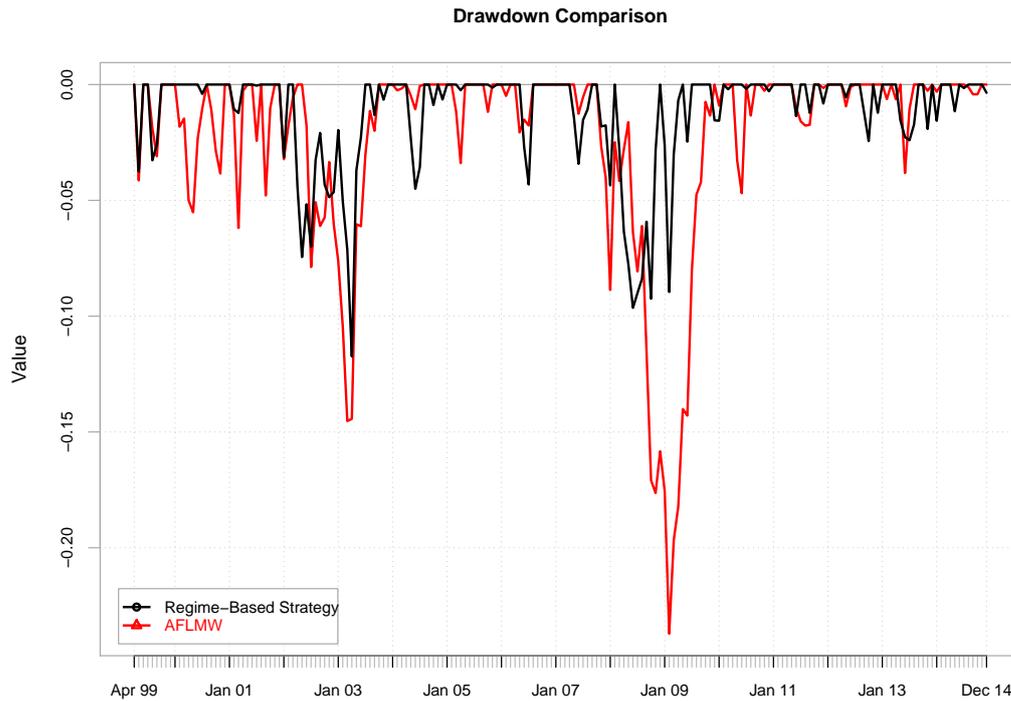
Drawdown Comparison

Figure 19: Comparison of Drawdowns from the different return series

The third and fourth moment in Table 14 provides the skewness and kurtoses measurements of the distributions. The skewness of a distribution measures the degree of asymmetry around the distribution mean (Cont, 2001). A positively skewed distribution implies that the asymmetric tail of the distribution lies more towards positive values. On the other hand, negative skewness implies a tail that extends more towards negative values of the distribution. From Table 14, it is observable that the regime-based strategy is positively skewed (1.76) while the AFLMW return index exhibits negative skewness (-0.0276). This implies that the regime-based strategy has more value that lie on the positive side of its distribution than the AFLMW index. In the next section, an analysis of the risk profile of the two return series.

6.2 ANALYSIS OF RISK

This section considers the risk metrics which are mostly used by practitioners in practice. The aim is to asses any observable "risk-reduction" features that were achieved through the regime-based strategy. In academic literature, numerous methods have been proposed as measures of portfolio risk. Perhaps the simplest and most well known of these measurements is the variance or standard deviation which was considered briefly in the previous section. However, as

mentioned in the previous section, one of the hindrances associated with variance as a measure of risk in practice was the fact that it considers an average of upside and downside deviations. This fails to capture the true risk of loss or rather downside risk which portfolio managers arguably equally concerned about. Furthermore, the use of standard deviation is mostly appropriate when the return distribution can be assumed to be normal. However from Table 14, we had observed that the distribution of our return series were skewed which nullifies the normality assumption.

In order to align this analysis for practical purposes, we consider risk metrics that are often used by practitioners. For our purpose, we consider six of these measurements, namely loss deviation, downside deviations, maximum drawdowns, historical VaR and historical Expected shortfall. We explain each of these in turn.

6.2.1 *Gains Deviation*

Even though this measure is not necessarily a risk measure per se, it has been included in this analysis in order to complement the analysis carried for the loss deviations and downside deviation. Informally, gains deviation is a measure of the average variability of positive returns in a portfolio (Satchel and Sortino, 2001). This measure captures favorable deviation in a return series that would translate into positive outperformance. This is calculated by taking the standard deviation of the positive returns of a series.

6.2.2 *Loss Deviations*

Loss deviation is a risk measure that measures the average variability of negative returns in a portfolio (Satchel and Sortino, 2001). This measures the downside risk in a portfolio that would result in capital losses. The computation of this measurement is similar to the computation of gains deviation but instead of calculating deviations from positive returns, negative returns are used.

6.2.3 *Downside Deviation*

Downside deviation is a risk measure similar to loss deviation but considers deviations in returns that fall below a minimum acceptable return (MAR) level (Satchel and Sortino, 2001). This risk measure is the same measure used in calculating the Sortino Ratio. Thus this risk measure seeks to capture the extent of potential impact from downside risk that would be translated into capital losses or negative deviation from a desired threshold.

6.2.4 *Value at Risk (VaR)*

In practice, VaR is one of the most widely used risk measure. Informally stated, VaR is a risk measure that measures the probability of extreme losses of an investment for a given period of time. In financial markets, most practitioner associate the VaR measure with possible investment losses that one would incur under normal market risk. This implies that when interpreting this measure, one needs to distinguish between normal and abnormal risk and between market and non-market risk. The modified or conditional VaR is an extension of the ordinary VaR which is calculated by taking a weighted average between VaR and any losses that are expected to exceed the specified VaR losses.

6.2.5 *Expected Shortfall*

Expected Shortfall risk measure is similar to that of the Conditional VaR measure. That is, Expected Shortfall measures the risk of incurring losses greater than those specified under the traditional VaR measure. The most notable difference between Conditional VaR and Expected Shortfalls is that Expected Shortfall does not require the assumption of normality in the return distribution.

6.2.6 *Analysis of the results*

Table 15 at the top of the next page provides a summary of the risk measures described above. In Table 15, the regime-based strategy resulted in a gains deviation of 2.56% while the AFLMW index resulted in a gains deviation of 2.15%. From these results, it can be argued that the proposed regime-based strategy could enhance upside potential by 0.41% higher than what managers are currently exposed to as observed from the AFLMW index. Also, the loss deviations resulting from the regime-based strategy (1.37%) were 0.425% better than the loss deviations observed in the AFLMW index (1.80%). Thus, accounting for regimes using the proposed model appears to provide protection against loss deviations relative to what managers are currently exposed to. These observations are supportive of the view of an attractive upside potential provided by the proposed regime-based strategy that was observed in the previous section and the downside risk reduction herein observed.

This point is further evident when considering downside deviation measures. From Table 15, tthe regime-based strategy exhibits downside deviations of 1.20% while the AFLMW index has downside deviation of 1.53%. The added downside protection that managers would obtain through the use of the regime-based strategy proposed by this thesis would be 0.33% on a monthly basis. Furthermore, the

Table 15: Analysis of risk

Risk Measure	Regime-Based Strategy	AFLMW Index
Gain Deviation	2.56%	2.15%
Loss Deviation	1.37%	1.80%
Downside Deviation	1.20%	1.53%
Maximum Drawdown	11.74%	23.71%
Historical VaR (95%)	−3.01%	−4.01%
Historical ES (95%)	−3.95%	−5.41%

maximum drawdown deferential between the regime-based strategy (11.74%) and the AFLMW index (23.71%) was 11.97%. This shows the drawdown benefit that could be achieved through efforts of trying to incorporate regime changes in the asset allocation decision.

Figure 20 on the next page provides the 95% Value at Risk and Conditional Value at Risk as modeled from the frequency distributions of the return series. From Figure 20, it can be observed that there are more return observations that lie within the 5% VaR tail in the AFLMW index than those observed from the proposed regime-based strategy. Similarly, from Table 15, it is observed that under the proposed regime-based strategy, the probability of experiencing extreme drawdown outcomes as measured by the historical VaR is 3.01% while the AFLMW index has a historical VaR of 4.01%. This implies that during extreme negative market moves, the proposed regime-based strategy would have exposed manages to less risk than their current historical exposure.

Complementing the findings from the VaR measurement, the results on Expected Shortfall measures also shows how accounting for regimes may potentially reduce the risk of drawdowns. From Table 15, it is observed that the Expected Shortfall of the regime-based strategy was 3.95% while the Expected Shortfall of the AFLMW index was observed to be 5.41%.

Thus, from these measures it can be argued that the proposed regime-based strategy exhibits attractive upside return potential with significant reduced downside risk. This shows that there are benefits that can be achieved through attempts of modeling regimes and incorporating them in asset allocation decisions. The attractiveness of the proposed model in this thesis is amplified by the ease at which macro-variables can be forecasted relative to financial variables (Pesaran, Schuermann and Smith, 2008). Findings from this section have provided some confidence on potential outperformance that can be achieved through extended efforts of regime classification in an asset allocation process and an anchor to the proposed regime model.
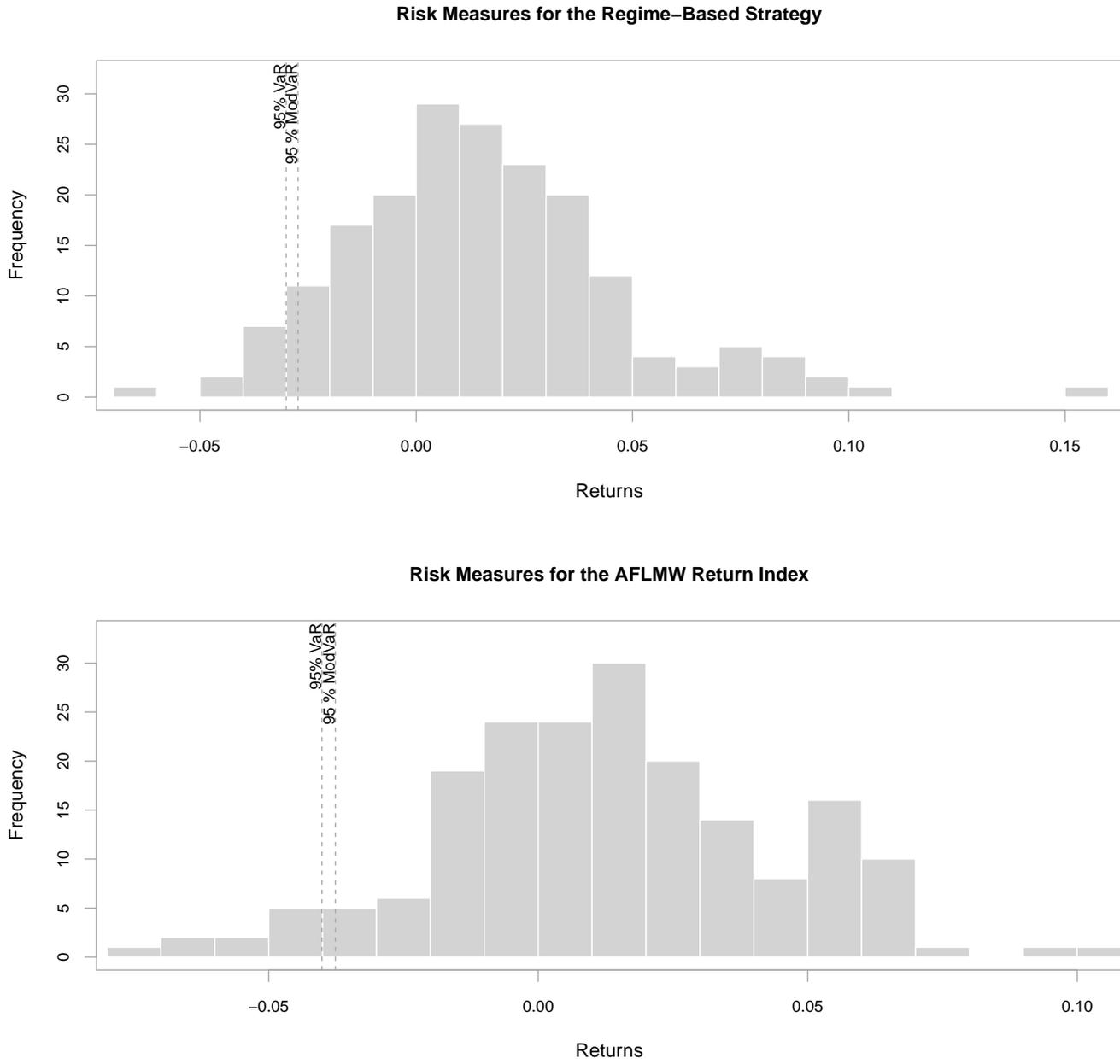
Risk Measures for the Regime–Based Strategy

Risk Measures for the AFLMW Return Index

Figure 20: VaR measures for the different return series

CONCLUSION

The main objective of this thesis as outlined in the thesis statement was twofold: firstly, we sought to model the cyclical nature of the investment environment through the construction of the so-called economic regimes. Secondly, the aim was to use the devised regime model to establish the extent of the potential advantages achievable by accounting for regime-dependent risk and return of asset classes.

In summary, we highlight key findings derived in answering the research question pertaining to this thesis. These can be summarised into four main points as follow.

1. *The employment of Principal Component Analysis (PCA) provided an attractive approach to exploring the multidimensional sphere of the real economy that characterises the cyclical nature of the financial markets. Thus, the major turbulent periods in the South African market were adequately captured by the classification model.*

   Empirical result established prior to the regime classification process illustrated the appropriateness of using Principal Components in modelling structural changes within an emerging economy. This approach was shown by Hun and Turner (2004) within a developed market environment to have provided significant information on the behaviour of the real economy in a reduced dimensional space. In particular, the authors demonstrated how the first and second PC's captured turbulent periods such as the Oil Shocks of the 1970's and 1980's, the savings loan crises of the early 1990's, the dot com bubble of the early 2000's and the credit crisis of 2008.

   In a similar vein, we have successfully demonstrated how using a set of macroeconomic-variables from the South African (SA) economy and applying Principal Component Analysis, we adequately capture market turbulence that affected the South African markets. More specifically, we have shown how the behaviour of the first principal component captured SA specific turbulence and those that were common to Emerging Markets. The most prominent market shocks that were captured by the first PC were the Asian crisis of 1998 - 1999, the Rand episode of the early 2000 and the global credit crisis of 2008. The appropriateness and adequacy of the PCA method established through these results provided a strong foundation and backbone for the regime classification processes.

83

2. *The novel approach of combining the PCA technique and Fuzzy Cluster Method provided a sensible classification of the different Economic Regimes which captured the time-varying risk return structure of Asset Classes.*

The novelty of this thesis and perhaps to some degree, a new thematic trail that requires further interrogation within the regime classification paradigm of asset allocation, is the combined application of PCA and Fuzzy Cluster Analysis. Indeed, to the best of our knowledge, no research was found that tackled the regime classification process in this manner. The novelty of this approach stem from two factors as highlighted within this thesis. Firstly, the PCA method provided a means of conceptualising the complexities of a multidimensional space that maps the real economy and financial markets. The idea behind the use of this method was not solely for dimension reduction but as a means of steering away from the elimination of variables when conducting our analysis.

Secondly and of equal importance was the use of the Fuzzy Clustering method. Inspired by the works of Hun and Turner (2004) and Dawsey (2014), this method constructed the different regimes cognisant of the "shades of grey" within each time period. That is, contrary to the hard clustering method followed by the aforementioned authors, the Fuzzy approach allows a more realistic and flexible approach to categorise time periods. The realism in the method was argued to lie in the fact that no single period purely belong or characterise a particular regime. Each time period, as given by the behaviour of the macro-variables, resemble each of the regimes to some degree. This provided a more interactive framework for investors to impose their subjective views of the world around them when determining which regime persist.

3. *The behaviour of the asset classes across the different Economic Regimes was consistent with empirical results from numerous authors under the subject.*

The different regimes that were constructed fully captured the time-varying nature of the different asset classes considered in this study. This was also consistent with the works of numerous authors that showed time-varying risk associated with financial assets (Yin and Yu Zhou, 2003; Guidolin and Timmermann, 2005 etc.). In particular, we demonstrated how no single asset was immune to all regimes. That is, similar to existing literature, we showed that under the defined regimes, no single asset class outperformed all other assets classes. Furthermore, we showed that the correlation structure between the asset classes

was different across the regimes. This result did not only show how South African markets exhibited time-varying risk but also provided support for the regime based model proposed in this thesis. Indeed, the ability to capture the time-varying nature of the different asset classes demonstrated by our regime classification model warranted its use. This further supported the use of the model in the asset allocation process that test portfolio performance based on regime classification.

4. *Accounting for Economic Regimes in the asset allocation process within an Emerging Market context provided an indication of the potential outperformance relative to strategies that were not regime cognizant.*

The portfolio optimisation process provided optimal portfolio compositions for each of the regimes. The allocation process in the optimizer was based on the perceived knowledge of the behaviour of the asset classes in each regime. Using the Centroid method proposed by Almgren and Chriss (2004), the optimisation was based on an ordinal ranking of the asset classes in descending order of performance in each regime. Thus it was evident from the plots of marginal contribution to portfolio risk that the ordinal structure was maintained within the allocation process. That is, top ranked assets in each of the regimes exhibited the most risk within the portfolios. The in-sample test of the strategy provided a stable and low risk portfolio which proved to outperform on average relative to strategies that were not regime cognizant. The resultant portfolio yielded an annualised return greater than that of the AFLMW return index demonstrating the potential that managers would achieve through the use of the proposed strategy. Furthermore, the portfolio yielded a very low risk profile of 2.1% with minimal deviations from this level over time. The upside potential of the strategy was shown to be very attractive with a maximum drawdown of 11.53% which was far less than the 23.43% observed from the AFLMW index. The average drawdown of the strategy was further shown to be less than 2% which signals a very stable portfolio.

7.1 FUTURE WORK

The analysis provided in this thesis interrogated an asset allocation strategy whose foundation was anchored on the modelling of the different regimes of the real economy. Although the results presented herein provided compelling support for the Regime-Based approach method proposed, it only provides yet another piece to the puzzling puzzle of portfolio construction. More importantly, as alluded to within

this work, our results kindled interesting questions under this thematic trait of Regime Based Asset allocation.

We elaborate on three key areas for future work pertaining the results presented in this thesis. These are summarised as follow.

1. Perhaps the greatest uncertainty relating to this study lies in the lack of an out-of-sample testing. However as argued in the outset of this study, there are two reasons that warrants the conclusions drawn without an out of sample test.

   Firstly, due to the major structural changes that occurred within the financial and real economy in South Africa, relevant data for analysis was limited. In total, the data under consideration covered a period of 20 years. In order to perform an out of sample testing, one would require an out of sample data set that would at least cover all defined regimes within the experimental or in sample data. Given the average length of each regime, in order to meet this criteria, one would expect a data set of at least 30 years. This was evidently not the case for South Africa.

   Secondly, in order to warrant an out-of-sample test, one would ordinarily expect to have performed an in-sample test that presented attractive performance features in a controlled environment. That is, a strategy that underperforms in an in sample test provides a lesser justification and need to test out of sample. This reason signifies the importance of an in-sample test as a starting point of an interrogative exercise.

   Hence, from these two points, a very critical area for further work is the performance of an out-of-sample test of the proposed strategy. Due to data limitations, numerous approaches can be undertaken in order to address an out of sample test. These may include the use of economic forecasts of the different time series, extrapolation and bootstrapping or simply undertaking a scenario study where asset returns are simulated.

2. It was also argued in this thesis that the novelty of this thesis lied in the proposed method for regime classification. Indeed, to the best of our knowledge, know literature was found that combined the two statistical techniques in the manner that was proposed in this study. However, even more importantly, within the area of Regime Based Asset Allocation, there is no standard approach proposed for the regime classification process. In particular, there are no set variables that seemed to be consistent across the literature for the regime classification methods. That is, there are no variables that have been tested for robustness in classifying different time period into regimes. This provides a huge area of research that is under researched.

This presents two points within this regards that requires further work. Firstly, due to time constraints and data limitations, the proposed method of using the PCA and Clustering Analysis method for regime classification was not tested for robustness. This is a critical point as any further improvements would require the method to be robust through time.

Secondly, an even more promising would be testing a set of variables that are best to use in the PCA decomposition for modelling the real economy. The excitement within this exercise lies in the fact that this has not yet been done. In currently existing literature, most researchers had used subjective reasoning on their choice of variables to use and the number of variables to use. However, no tested method has been provided that tests for the significance of the commonly accepted variables. More importantly, the number of variables required to capture the behaviour of the real economy was not fully interrogated.

3. Lastly, in this study, we considered an asset allocation strategy on an asset class level. Further work could consider a similar analysis but done on a sector level. In this case, the aim would be to test the performance of different sectors under the different regimes. This would be valuable since it would give a ranking of sectors where stock picking would occur for each asset class

Furthermore, one can also consider different investment styles under the different regimes. That is, it would also be interesting to interrogate investment strategies such as momentum style, value style, Growth, Quality or High volatility investment styles.

Part IV

APPENDIX

APPENDIX TEST

The table below lists the economic indicators that were used by Hun and Turner (2010)

Table 16: Economic indicators - Source: St. Louis Fed

Name	Description
umcsent	University of Michigan: Consumer Sentiment
umcsent3d	3-month Change of umcsent
lgdpc12d	12-month Change in log of GDP
lgdpc3d	3-month Change in log of GDP
tdsp	Household Debt Service Payments
tdsp3d	3-month Change of tdsp
lpce3d	3-month Change in Log of Personal Cons. Exp.
lpce logdpc1	log(pce) - log(gdpc1)
unrate	Civilian Unemployment Rate
unrate3d	3-month Change of Unemployment Rate
payems1d	1-month Change of Nonfarm Payrolls
credit	Credit Spread (Baa - Aaa)
term	30-Year Treasury- 10-Year Treasury CM Rate
dfedtar	Federal Funds Target Rate
gs10	10-Year Treasury Constant Maturity Rate
oilprice	Oil Price - West Texas Intermediate
cpifesl12d	CPI All Items Ex Food & Energy - 12-month Change
dgorder12d	Durable Goods Order - 12-month Change
dgorder3d	Durable Goods Order - 3-month Change
napm	ISM Manufacturing: PMI Composite Index
tcu	Capacity Utilization: Total Industry
lindpro12d	Industrial Production Index
lipmat12d	Industrial Production: Materials
liputil12d	Industrial Production: Electric and Gas Utilities
lipmine12d	Industrial Production: Mining
lipbuseq12d	Industrial Production: Business Equipment
lipcongd12d	Industrial Production: Consumer Goods

BIBLIOGRAPHY

[1] Alemgren, R. and Chriss, N. (2004) Optimal Portfolios from Ordering Information, available: http://www.courant.nyu.edu/~almgren/papers/optport.pdf.

[2] Amenc, N., Malaise, P., Martellini, L., and Sfeir, D. (2003) Tactical Stle Allocation A New Form of Market Neutral Strategy, available: http://docs.edhec-risk.com/EHFD-2006B/doc/Tactical20Style20Allocation.pdf.

[3] Ammann, M. and Verhofen, M. (2005) The Effect of Market Regimes on Style Allocation, available: http://www.fbv.kit.edu/symposium/10th/papers/Verhofen20-20The20Effect20of20Market20Regimes20on20Style20Allocation.pdf.

[4] Araujo, M. and Costa, O. (2008) Multi period Mean Variance Portfolio Optimisation With Markov Switching Parameters, available: http://www.scielo.br/pdf/ca/v19n2/a03v19n2.

[5] Aug, A. and Bekaert, G. (2002) How do Regimes Affect Asset Allocation, available: https://www0.gsb.columbia.edu/faculty/aang/papers/inquire.pdf.

[6] Auret, C. and Golding, J. (2012) Stock prices as a leading indicator of economic activity in South Africa: Evidence from the JSE, available: http://www.iassa.co.za/wp-content/uploads/IAJ76-No4-Auret-Golding-final.pdf.

[7] Auret, C. and Vivian, R. (2014) A comparative analysis of returns of various financial asset classes in South Africa: a triumph of bonds, available: http://www.unisa.ac.za/contents/faculties/service_dept/docs/Sabview_18_3_Chap208.pdf.

[8] Avouyi-Dovi, S. and Matheron, J. (2003) Interaction between business cycles financial cycles and monetary policy Stylised facts, available: http://www.bis.org/publ/bppdf/bispap22p.pdf.

[9] Avouyi-Dovi, S. and Matheron, J. (2004) Interactions between business cycles, financial cycles and monetary policy: stylised facts1, available: http://www.bis.org/publ/bppdf/bispap22p.pdf.

[10] Bae, G., Kim, W., and Mulvey, J. (2013) Dynamic asset allocation for varied financial markets under regime switching framework, available: http://www.sandpointe.com/pdf/

dynamic-asset-allocation-for-regime-switching-framework.
pdf.pdf.

[11] Balkissoon, K., Bennett, R., Carl, P., Lestel, M., Peterson, B.,
Ulrich, J., Wuertz, D., and Zivot, E. (2014) PerformanceAn-
alytics, available: https://cran.r-project.org/web/packages/
PerformanceAnalytics/PerformanceAnalytics.pdf.

[12] Banz, R. (1981) The relationship between return and market
value of common stock, available: http://www.business.unr.
edu/faculty/liuc/files/BADM742/Banz_sizeeffect_1980.
pdf.

[13] Besnainou, I., Jordan, J., and Portiat, R. (2001) Dynamic Asset Al-
location for Stocks Bonds and Cash, available: http://www.nera.
com/content/dam/nera/publications/archive1/3671.pdf.

[14] Bevan, A. and Winkelmann, K. (1998) Using the Black-Litterman
Global Asset Allocation Model: Three Years of Practical Expe-
rience, available: https://faculty.fuqua.duke.edu/~charvey/
Teaching/IntesaBci_2001/GS_Using_the_black.pdf.

[15] Bezdek, J., Ehrich, R., and Full, W. (1984) FCM:
The Fuzzy c-Means Clustering Algorithm, avail-
able: http://www.cse.hcmut.edu.vn/~chauvtn/
data_mining/Reading/Chapter20520-20Clustering/
FCM20-20The20Fuzzy20c-Means20Clustering20Algorithm.pdf.

[16] Biglova, A., Fabozzi, F., Rachev, S., and Rachev, S. (2004)
An Empirical Examination of Daily Stock Return Distribution
for US Stocks, available: http://www.ams.sunysb.edu/~rachev/
publication/tr_an_empirical_examination.pdf.

[17] Black, F. and Litterman, R. (1992) Global Portfolio Optimi-
sation, available: http://www.sef.hku.hk/tpg/econ6017/2011/
black-litterman-1992.pdf.

[18] Bleaney, M., Misen, P., and Veleanu, V. (2012) Credit Spreads as
Predictors of Economic Activity in Eight European Economies,
available: https://editorialexpress.com/cgi-bin/
conference/download.cgi?db_name=MMF2012&paper_id=158.

[19] Blitz, D. and van Vleit, P. (2009) Dynamic Strategic As-
set Allocation Risk and Return across Economic Regimes,
available: http://www.master272.com/finance/memoire_2014/
strategies/sector20rotation/dynamic20SAA.pdf.

[20] Bodington, L. (2014) Gold in South African market A safe
haven or hedge, available: http://wiredspace.wits.ac.za/
bitstream/handle/10539/16822/RESEARCH20REPORT.pdf?
sequence=1.

[21] Borio, C. (2012) The Financial Cycle and macroeconomics What have we learnt, available: http://www.bis.org/publ/work395.pdf.

[22] Bourachnikova, O. and Yusupov, N. (2011) Symmetric vs. Downside Risk Measures in Portfolio Decisions, available: http://www.revue-banque.fr/medias/content/users/christine/1315320017991.pdf.

[23] Briere, M. and Signori, O. (2011) Inflation hedging portfolios in different regimes, available: http://www.bis.org/publ/bppdf/bispap58g.pdf.

[24] Bulla, J., Mergner, S., Sesboue, A., and Chesneeau, C. (2010) Markov switching Asset Allocation Do Profitable Strategies Exist, available: http://mpra.ub.uni-muenchen.de/21154/1/MPRA_paper_21154.pdf.

[25] Burio, C. (2012) The financial cycle and macroeconomics: What have we learnt?, available: http://www.bis.org/publ/work395.pdf.

[26] Cannon, R., Dave, J., and Bezdek, J. (1986) Efficient Implementation of the Fuzzy c-Means Clustering Algorithms, available: http://cb.uu.se/~joakim/course/fuzzy/vt10/CE1/Bezdek_FCM.pdf.

[27] Chad, S. and Giles, E. (2003) Income Convergence and Trade Openness: Fuzzy Clustering and Time Series Evidence, available: http://www.uvic.ca/socialsciences/economics/assets/docs/econometrics/ewp0304.pdf.

[28] Chalmers, J., Kaul, A., and Phillips, B. (2011) The Wisdom of Crowds Mutual Fund Investors Aggregate Asset Allocation Decisions, available: http://www.researchgate.net/publication/228203560_The_Wisdom_of_Crowds_Mutual_Fund_Investors'_Aggregate_Asset_Allocation_Decisions.

[29] Chan, L. and Lakonishok, J. (2004) Value and Growth Investing: Review and Update, available: http://lsvasset.com/pdf/research-papers/Value_and_Growth_Investing_FINAL.pdf.

[30] Chin, C. (2013) Correlations Have Personality Too An Analysis of Correlations between Assets, available: http://www.naaim.org/wp-content/uploads/2013/10/00L_Correlations_have_Personality_carlton_chin.pdf.

[31] Claessens, S., Kose, M., and Terrones, M. (2011) Financial Cycles: What? How? When?, available: https://www.imf.org/external/pubs/ft/wp/2011/wp1176.pdf.

[32] Classens, S., Kose, M., and Terromes, M. (2011) How do Business and Financial Cycles Interact?, available: https://www.imf.org/external/pubs/ft/wp/2011/wp1188.pdf.

[33] Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues, available: http://www.proba.jussieu.fr/pageperso/ramacont/papers/empirical.pdf.

[34] Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues., available: http://www.proba.jussieu.fr/pageperso/ramacont/papers/empirical.pdf.

[35] Davis, J., Aliaga-DÃaz, R., and Patterson, A. (2011) Asset allocation in a low-yield and volatile environment, available: https://personal.vanguard.com/pdf/s286.pdf.

[36] Dawsey, K. Unpublished (2014) US Daily: Where Are We in the Cycle?, available: http://www.abnamromarkets.com/fileadmin/user_upload/TA/2014/US_Daily__Where_Are_We_in_the_Cycle___Dawsey_.pdf.

[37] Dubey, S. and Ghosh, S. (2013) Comparative Analysis of K-Means and Fuzzy C-Means Algorithms, available: https://thesai.org/Downloads/Volume4No4/Paper_6-Comparative_Analysis_of_K-Means_and_Fuzzy_C_Means_Algorithms.pdf.

[38] Duda, R., Hart, P., and Stock, D. (1995) Pattern Classification and Scene Analysis, available: http://www.svms.org/classification/DuHS95.pdf.

[39] Duncan, A. and Kabundi, A. (2011) Volatility Spillovers across South African Asset Classes during Domestic and Foreign Financial Crises, available: http://www.econrsa.org/w_papers/wp202.pdf.

[40] Emsbo-Mattingly, L., Hofschire, D., Litvak, A., and Lund-Wilde, J. (2014) A Business Cycle Approach to Equity Sector Investing, available: https://scs.fidelity.com/common/application/markets_sectors/business_cycle/Business_Cycle_Sector_Approach.pdf.

[41] Etula, E. and Farshid, M. (2012) Strategic Asset Allocation in a Multi Factor World, available: www.iinews.com/site/pdfs/gs_fall_2012_journal_pm.pdf.

[42] Eun, C., Kilc, R., and Lai, S. (2012) A Tale of Two Exchange Rates: South Africa Dual-Rate Experiment, available: https://webcache.googleusercontent.com/search?q=cache:jRk7OVtEAaAJ:https://www.frbatlanta.org/~/media/Documents/research/seminars/2012/seminarkilic060112.pdf+&cd=1&hl=en&ct=clnk&gl=za.

[43] Eychenne, K., Martinetti, S., and Rocalli, T. (2011) Strategic Asset Allocation., available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2154021.

[44] Fama, E. and French, K. (1992) The Cross-Section of Expected Stock Returns, available: http://www.bengrahaminvesting.ca/Research/Papers/French/The_Cross-Section_of_Expected_Stock_Returns.pdf.

[45] Faulkner, D. and Loewald, C. (2014) Policy Changes and Economic Growth A Case Study of South Africa, available: http://www.econrsa.org/system/files/publications/policy_papers/pp14.pdf.

[46] Forest, N., Orpiszewski, T., A., P., and Kojo, G. (2014) RIDING THE CYCLES: Understanding business, financial and monetary cycles in order to allocate fixed income, available: https://www.candriam.com/siteassets/medias/publications/whitepapers/can_whitetotalreturn_eng.pdf.

[47] Gilchrist, S. and Zakrajsek, E. (2011) Credit Spreads and Business Cycle Fluctuations, available: http://people.bu.edu/sgilchri/research/NBER_w17021.pdf.

[48] Glennie, R., Kneul, T., Langrock, R., and Michelot, T. (2015) Markov switching generalised additive models, available: http://arxiv.org/pdf/1406.3774.pdf.

[49] Godfeld, S. and R.E, Q. (1973) The Estimation of Structural Shifts by Switching Regression, available: http://www.nber.org/chapters/c9938.pdf.

[50] Graflund, A. and Nilsson, B. (2001) Dynamic Portfolio Selection The Relevance of Switching Regimes and Investment Horizon, available: http://stat-arb.org/pdf/HMM/Graflund.pdf.

[51] Grover, N. (2014) A study of various Fuzzy Clustering Algorithms, available: http://www.ijer.in/ijer/publication/v3s3/IJER_2014_310.pdf.

[52] Guidolin, M. and Timmermann, A. (2005) Optimal Portfolio Choice under Regime Switch Skew and Kurtosis Preferences, available: https://research.stlouisfed.org/wp/2005/2005-006.pdf.

[53] Hamilton, J. (1989) A New Approach to the Economic Analysis of Non-stationary time series and the business cycle, available: http://www.ssc.wisc.edu/~bhansen/718/Hamilton1989.pdf.

[54] Hanival, S. and Maia, J. (2012) An Overview of the Performance of the South African Economy since 1994, available:

http://www.thepresidency.gov.za/docs/reports/15year_
review/economic/economic_performance.pdf.

[55] Han, J. and Turner, C. (2010) Optimization under Time-
Varying Economic Regimes, available: http://cs229.stanford.
edu/proj2009/HanTurner.pdf.

[56] He, G. and Litterman, R. (1999) The Intuition behind Black-
Litterman Model Portfolio, available: http://papers.ssrn.com/
sol3/papers.cfm?abstract_id=334304.

[57] Herold, U. (2003) Portfolio Construction with Qualita-
tive Forecasts, available: http://www.cfapubs.org/doi/pdf/10.
2469/dig.v34.n2.1429.

[58] Holland, S. (2006) CLUSTER ANALYSIS, available: http://
strata.uga.edu/software/pdf/clusterTutorial.pdf.

[59] Hung, M. and Yang, D. (2001) An efficient
Fuzzy C-Means Clustering Algorithm, available:
http://www.csee.usf.edu/~manohar/Papers/FCM/
An20efficient20Fuzzy20C20Means20clustering20Algorithm.
pdf.

[60] Islam, I. and Verick, S. (2010) The Great Recession of 2008-2009:
Causes, Consequences and Policy Response, available: http://
ftp.iza.org/dp4934.pdf.

[61] Jolliffe, I. (2002) Principal Component Analysis, available: http:
//wpage.unina.it/cafiero/books/pc.pdf.

[62] Kannan, P., Scott, A., and Terrones, M. (2009) From Re-
cession to Recovery: How Soon and How Strong, avail-
able: https://www.imf.org/external/np/seminars/eng/2012/
fincrises/pdf/ch8.pdf.

[63] Kaplan, P. (1998) Asset Allocation Models Using the
Markowitz Approach, available: http://www.blacklitterman.
org/Black-Litterman.pdf.

[64] Kneib, T., Langrock, R., and Sohn, A. (2015) Non-parametric
Hidden Markov Models, available: http://www.rolandlangrock.
com/talks/SCo.pdf.

[65] Kollar, M. (2013) A Sketch of Macro Based Asset Al-
location, available: http://www.iises.net/download/Soubory/
soubory-puvodni/pp101-120_ijoes_2013V2N3.pdf.

[66] Ledoit, O. and Wolf, M. (2003) Honey I Shrunk the Sample Co-
variance Matrix, available: http://www.ledoit.net/honey.pdf.

[67] Levy, P. (1999) Sanctions on South Africa What did they do, available: `http://www.econ.yale.edu/growth_pdf/cdp796.pdf`.

[68] Lewis, J. (2001) Policies to Promote Growth and Employment in South Africa, available: `http://adapt.it/adapt-indice-a-z/wp-content/uploads/2015/02/lewis_2001.pdf`.

[69] Lyngby, K. (2014) Regime-Based Asset Allocation: Do Profitable Strategies Exist, available: `http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6808/pdf/imm6808.pdf`.

[70] Madhulatha, T. (2012) An Overview on Clustering Methods, available: `http://arxiv.org/ftp/arxiv/papers/1205/1205.1117.pdf`.

[71] Markowitz, H. (1952) Portfolio Selection, available: `https://www.math.ust.hk/~maykwok/courses/ma362/07F/markowitz_JF.pdf`.

[72] Masih, R. (2011) Decision Making Under Different Economic Regimes., available: `https://www.jpmorgan.com/jpmpdf/1320567053393.pdf`.

[73] Merton, R. (1969) Lifetime Portfolio Selection Under Uncertainty The Continuous Time Case, available: `http://www.people.hbs.edu/rmerton/Lifetime20Portfolio20Selection.pdf`.

[74] Mueller, P. (2007) Credit Spreads and Real Activity, available: `http://www.lse.ac.uk/fmg/documents/events/seminars/lunchtime/993_PhMueller.pdf`.

[75] Munro, B. and Silberman, K. Unpublished (2008) Optimal asset allocation in different economic environments., available: .

[76] Pask, A. (2008) South African Asset Classes Return and Volatility relationship dynamics over time, available: `http://uir.unisa.ac.za/bitstream/handle/10500/2707/dissertation_pask_20a.pdf?sequence=1`.

[77] Pesaran, M., Schuermann, T., and Smith, L. (2008) Forecasting Economic and Financial Variables with Global VARs, available: `https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr317.pdf`.

[78] Pretorius, W. and Venter, J. (2004) Note on the revision of composite leading and coincident business cycle indicator, available: `https://www.resbank.co.za/Lists/News20and20Publications/Attachments/4411/Note20on20the20revision20of20composite20leading20and20coincident20business20cycle20in... pdf`.

[79] Pule, B. and Mokoaleli, T. (2013) Evaluation of Gold as an Investment Asset The South African context, available: `http://wiredspace.wits.ac.za/bitstream/handle/10539/12911/Evaluation20of20Gold20as20an20Investment20Asset_The20South20African20Context_BP20Pule.pdf?sequence=1`.

[80] Sa-Aadu, J., Shilling, J., and Tiwari, A. (2005) Portfolio Performance and Strategic Asset Allocation Across Different Economic Conditions, available: `http://www.biz.uiowa.edu/faculty/atiwari/Diversification.pdf`.

[81] Satchell, S. and Scowcroft, A. (2000) A demystification of the Black-Litterman model: Managing quantitative and traditional portfolio construction, available: `http://www.palgrave-journals.com/jam/journal/v1/n2/abs/2240011a.html`.

[82] Satchel, A. and Sortino, F. (2001) Managing downside risk in financial markets: Theory, Practice and Implementation., available: `http://elis.dvo.ru/~lab_11/bib/Sortino.pdf`.

[83] Sheikh, A. and Sun, J. (2011) Regime Changes: Implications of macro shifts on asset class and portfolio performance, available: `https://www.jpmorgan.com/jpmpdf/1158630464604.pdf`.

[84] Shlens, J. (2003) A Tutorial on Principal Component Analysis, available: `https://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf`.

[85] Sortino, F. and Price, L. (1994) Performance Measurement in a Downside Risk Framework, available: `http://www.iijournals.com/doi/abs/10.3905/joi.3.3.59?journalCode=joi`.

[86] Steinbach, M. (1999) Single Period and Multi Period Mean Variance Models, available: `file:///C:/Users/mpumelelok/Downloads/SC-99-30.pdf`.

[87] Van der Merwe, E. (2004) Inflation targeting in South Africa, available: `https://www.resbank.co.za/Lists/News20and20Publications/Attachments/4704/OCCNo19[1].pdf`.

[88] Walt, B. (1982) Indicators of business cycle changes in South Africa, available: `https://www.resbank.co.za/Lists/News20and20Publications/Attachments/5248/07Indicators20of20business20cycle20changes20in20South20Africa.pdf`.

[89] Walter, J. (2007) The Black Litterman Model in Detail, available: `http://www.blacklitterman.org/Black-Litterman.pdf`.

[90] Yin, G. and Yu Zhou, X. (2003) Markowitz Mean Variance Portfolio Selection with Regime Switching From Discrete time Models to their Continuous time limits, available: http://people.maths.ox.ac.uk/~zhouxy/download/mvjump_part2.pdf.

[91] Yin, G. and Yu Zhou, X. (2006) Markowitz Mean Variance Portfolio Selection with Regime Switching A Continuous Time Model, available: http://people.maths.ox.ac.uk/~zhouxy/download/mvjump_part1_rev.pdf.

Name	Mpumelelo Kondlo
Student No:	KNDMPU001
Tel numbers:	0783652492
Word count:	-
Dissertation Title:	A Framework for Regime Identification and Asset Allocation
Name of Supervisor/s:	Professor David Bradfield
Email address:	mpumelelo.kondlo@gmail.com
No. of pages	115

Declaration:

1. I am presenting this dissertation in 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.

Cape Town, March 2016

By: Mpumelelo Kondlo