Title	Applications of machine learning in finance: analysis of international portfolio flows using regime-switching models
Author(s)	Ó Cinnéide, Ruairí
Publication date	2019
Original citation	Ó Cinnéide, R. 2019. Applications of machine learning in finance: analysis of international portfolio flows using regime-switching models. MRes Thesis, University College Cork.
Type of publication	Masters thesis (Research)
Rights	**© 2019, Ruairí Ó Cinnéide.** https://creativecommons.org/licenses/by-nc-nd/4.0/
Item downloaded from	http://hdl.handle.net/10468/10895

Ollscoil na hÉireann, Corcaigh

National University of Ireland, Cork

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Applications of Machine Learning in Finance: Analysis of International Portfolio Flows using Regime-Switching Models

Thesis presented by

Ruairí Ó Cinnéide, BEng

For the degree of MSc (Commerce)

University College Cork

Department of Business Information Systems

Head of Department: Professor Joseph Feller

Supervisors: Dr. John O'Brien, Prof. Mark Hutchinson

2019

Table of Contents

Acknowledgements 5

Abstract 6

Chapter 1. Introduction 7

Chapter 2. Machine Learning in Quantitative Finance 11

2.1. Introduction 11

2.2. Background 13

2.2.1. Machine Learning 13

2.2.2. Common Machine Learning Approaches and Algorithms 13

2.2.3. The Evolution of Quantitative Investing 15

2.3. Methodology 17

2.4. Results 18

2.4.1. Popular Machine Learning Use Cases and Algorithms 18

2.4.2. Summary of Key Insights from Recent Papers 20

A. Portfolio Construction 20

B. Return Forecasting 23

C. Risk 26

2.5. Discussion 29

2.5.1. Strategy Development and Analysis 29

2.5.2. The Use of Alternative Data 30

2.5.3. Choosing Machine Learning Algorithms 30

2.5.4. Backtesting and Strategy Verification 31

2.6. Conclusion 32

Chapter 3. Regime Switching Models 33

3.1. Introduction 33

3.2. Capturing Stylized Facts using HMMs 34

3.3. Regime Switching in Asset Allocation 34

3.4. Model Selection 35

3.5. Hidden Markov Models 37

Chapter 4. A Flow-based Regime-Switching Model: Data & Methodology 39

4.1. Introduction 39

4.2. International Portfolio Flows 40

4.3. Flow Data 40

4.4. Price Data 42

4.5. Training & Characterising Regimes using Flow data 44

4.5.1. Persistence of Flows vs Log Returns 44

4.5.2. Identifying & Characterising Regimes present in Flow Data 45

4.5.3. Model Selection 45

4.6. A Regime-Based Asset Allocation Strategy 46

Chapter 5. A Flow-based Regime-Switching Model: Results 49

5.1. Training & Characterising Regimes using Flow Data 49
5.1.1. Persistence of Flows vs Returns 49
A. Daily Frequency 50
B. Weekly Frequency 51
5.1.2. Model Switching Behaviour & Structural Breaks 52
5.1.3. Characteristics of Regimes 58
A. Daily Frequency 58
B. Weekly Frequency 59
5.1.4. Model Selection 61
5.2. Testing a Regime-Based Asset Allocation Strategy 62
5.2.1. Daily Frequency 62
5.2.2. Weekly Frequency 66
5.2.3. Robustness Tests 68
Chapter 6. Conclusion 70
Bibliography 73
Appendix 1: Model Switching Behaviour 83
Appendix 2: ICEFR 2019 Conference Paper 90
Appendix 3: CFA European Quant Awards 2019 - 2nd Place Finish 1

Declaration

The author declares that except where duly acknowledged, this thesis is entirely his own work and has not submitted for any degree in the National University of Ireland, Cork or any other University.

Acknowledgements

Firstly, I would like to thank my supervisor Dr. John O'Brien for his support throughout the course of my research and the completion of my thesis. He made himself fully available for advice and guidance in person or through email if I came to him with any questions.

I would like to thank Michael Metcalfe and Marija Veitname from State Street Global Markets for the insights and advice they provided over countless phone calls relating to the research. I would also like to thank Brian McCabe and Neasa Ní Fhathartha from State Street Corporation for their support through the excellent facilities provided in the UCC State Street Advanced Technology Centre research lab, and for aiding in the overall research process.

I would like to thank my research partners, Sophie Emerson and Luke O'Shea for their support and collaboration, I have no doubt that much of our progress wouldn't have been achieved without our good teamwork. I would also like to thank my other lab-mates in the State Street Advanced Technology Centre for all the laughs and discussions we had during the year.

Finally, I would like to thank my parents Ann and Thos, my sister Róisín, my girlfriend Grace, and my friends for all their support and encouragement while completing my thesis and in all other aspects of life.

Ruairí Ó Cinnéide

Abstract

Recent advances in machine learning are finding commercial applications across many sectors, not least the financial industry. This thesis explores applications of machine learning in quantitative finance through two approaches.

The current state of the art is evaluated through an extensive review of recent quantitative finance literature. Themes and technologies are identified and classified, and the key use cases highlighted from the emerging literature. Machine learning is found to enable deeper analysis of financial data and the modelling of complex non-linear relationships within data. The ability to incorporate alternative data in the investment process is also enabled. Innovations in backtesting and performance metrics are also made possible through the application of machine learning.

Demonstrating a practical application of machine learning in quantitative finance, regime-switching models are applied to analyse and extract information from international portfolio flows. Regime-switching models capture properties of international portfolio flows previously found in the literature, such as persistence in flows compared to returns, and a relationship between flows and returns. Structural breaks and persistent regime shifts in investor behaviour are identified by the models. Regime-switching models infer regimes in the data which exhibit unique characteristic flows and returns.

To determine whether the information extracted could aid in the investment process, a portfolio of global assets was constructed, with positions determined using a flow-based regime-switching model. The portfolio outperforms two benchmarks, a buy & hold strategy and the MSCI World Index in walk-forward out-of-sample tests using daily and weekly data.

Chapter 1. Introduction

The material presented in this thesis is based on twelve months of research conducted in the State Street Advanced Technology Centre, an industry-focused academic research collaboration between State Street Corporation and University College Cork.

In recent years, there has been increased interest in the use of machine learning across a variety of sectors, with finance being no exception. While machine learning is being adopted across the financial industry as a whole, the focus of this research is on applications of machine learning within investment management.

There were two main objectives motivating this research. The first was to explore current and potential applications of machine learning to the investment process as the field has grown and developed. In particular, this includes the development of machine learning applications for return forecasting, portfolio construction and risk modelling. The second was to extract investment information from international portfolio flow data provided by State Street Global Markets, through the application of machine learning.

To accomplish the first objective, a review was conducted of the growing body of academic literature applying machine learning to investment problems. The approach taken was to first develop a broad understanding of machine learning by surveying the main frameworks, techniques and programming languages used, as well as examining historically how investors have adapted to shifts in technology, tools and methodologies. After this, an extensive literature review was carried out of the latest research papers applying machine learning across the areas of return forecasting, portfolio construction and risk modelling. The results of this literature review provided key information for a discussion of current trends into the integration and adoption of machine learning by investors, as well as providing a good grounding when approaching the second research objective.

International portfolio flows describe the actions of informed investors, shareholders, and fund managers who add or remove cash from funds and buy or sell individual securities with fund deposits. Flows have been shown to be stable and persistent in nature. They have also been shown to influence equity returns, and can be used to measure market sentiment, improve the timing of asset class specific risk, and inform asset allocation strategies. The second research objective was to determine if it was possible to extract information from portfolio flow data using machine learning, and that see if that information could form the basis of a proof-of-concept investment strategy, constructed with insight from State Street Global Markets. State Street Global Markets conduct research into international portfolio flows constructed using proprietary capital allocation data from thousands of institutional investor portfolios under custody and administration by State Street Corporation. They also offer research, trading and securities lending services across foreign exchange, stocks, fixed income and derivatives. As well as providing the portfolio flow data, they also provided practical industry insight and feedback throughout the research and when developing the proof-of-concept.

It was decided to analyse the flow data using a regime-switching model. Regimes are periods of time with unique characteristic financial variables. Regime-switching models make use of Hidden Markov models, an unsupervised machine learning technique, to identify regimes in financial data. Regime-switching and Hidden Markov models have been successful in finance literature across a number of asset-allocation strategies, with the majority of the literature focusing on the analysis of returns. Combining regime-switching models and portfolio flow data provided an interesting opportunity to apply a machine learning technique with established success in finance to a dataset with interesting properties, in a specific combination which had not previously occurred in the literature.

The thesis is structured as follows. Chapter 2 contains a literature review and discussion of the trends and applications of machine learning in quantitative finance. Chapter 3 introduces the concept of regime-switching models, detailing their applications in the literature, and describing their functionality. Chapter 4 discusses international portfolio flows, previous research into the properties of flows, as well as discussing the specific State Street Corporation dataset used in this research. The methodology applied when analysing the portfolio flow data using regime-switching models is also discussed, as well as the approach taken when constructing a regime-based asset allocation strategy. Chapter 5 contains the results obtained modelling and characterising the portfolio flow data using regime-switching models, as well as the performance of a portfolio of global indices managed using a regime-based asset allocation strategy. Chapter 6 provides a conclusion and summary of the entire body of research, with a discussion of the results achieved towards both research objectives. Appendix 1 contains additional figures and discussion relating to the behaviour of regime-switching models in Chapter 5. Appendix 2 contains a conference paper "*Trends and Applications of Machine Learning in Quantitative Finance*" presented at ICEFR 2019, which formed the basis for Chapter 2. Appendix 3 contains a research report summarizing the results of Chapters 3-5, "*Dynamic Regime-Based Asset Allocation using International Equity Flow Data*". This report achieved 2nd place at the CFA European Quant Awards 2019.

Both research objectives were formed as part of a three-person research team (including the author, Sophie Emerson, Luke O'Shea, supervised by Dr. John O'Brien) in the State Street Advanced Technology Centre, focusing on applications of machine learning in finance.

When accomplishing the first research objective, covered in Chapter 2 of the thesis, the research methodology was devised and agreed upon by the entire research team. When reviewing background literature (Section 2.2) it was decided to separately discuss the topics of machine learning and finance. The author was responsible for conducting a review on the evolution of quantitative finance.

The task of conducting a review of the latest academic literature was divided equally between the researchers. The results of this review (Section 2.4) were compiled by the whole research team, working in tandem to create a concept-centric matrix and identifying recurring themes in the literature. The research team also worked in tandem to summarise a large body of literature across the various areas of finance where machine learning is being applied. The author was responsible for the majority of the discussion of the results of the literature review (Section 2.5), and for the editing of the introduction and conclusion of Chapter 2.

The second research objective, determining whether it was possible to extract information from portfolio flow data using machine learning, and using that information as the basis for a proof-of-concept investment strategy, was decided upon by the research team, with input from State Street Global Markets. After this research objective was decided upon, each member of the research team conducted individual research into potential machine learning approaches for extracting information from portfolio flow data. Chapters 3-5 of the thesis contain solely the authors approach to this research objective, which utilized regime-switching models.

The introduction and conclusion of the thesis were written by the author.

Chapter 2. Machine Learning in Quantitative Finance

2.1. Introduction

Machine Learning is a subfield of Artificial Intelligence (AI) that uses statistical techniques that provide computer models with the ability to learn from a dataset, allowing the models to perform specific tasks without explicit programming (Domingos 2012). Machine learning is being applied to improve function across the finance industry in a wide range of areas including, for example, fraud detection, payment processing and regulation. This chapter evaluates current and potential applications of machine learning to the investment process. In particular, this includes the development of machine learning applications for return forecasting, portfolio construction and risk modelling.

The first widespread commercial use cases of artificial intelligence were "expert systems", originating in Stanford in the 1960s (Lindsay, Buchanan et al. 1993) and popularised in the 1980s and 1990s. Expert systems were designed to solve complex problems in a specific field, in a manner similar to a subject matter expert. Original expert systems were rule-based programs developed in languages such as LISP and Prolog. In recent years, there has been a significant drop in interest in classic expert systems, as they are superseded by systems incorporating artificial intelligence (Wagner 2017). AI systems are systems that replicate human thought processes. (Russell and Norvig 2009). Many of these systems are advertised today as cognitive computing systems.

Cognitive computing describes a computer system which mimics human cognitive process in some way, cognitive processes are those that allow individuals to remember, think, learn and adapt (Modha and Witchalls 2014). The term has gained recognition in the public domain in recent years, due in large to the introduction of Watson, IBM's cognitive computing system. These systems are constructed by combining computer science with statistical and machine learning techniques developed over the last century (Domingos 2012). Watson, in its original form, was a question answering computing system, responding to questions posed in natural language. It was introduced on the television quiz show "Jeopardy!" – where it defeated two of the show's most celebrated contestants in the "IBM Challenge" (Ferrucci, Brown et al. 2010). Large-scale systems such as Watson combine many techniques (Ferrucci, Brown et al. 2010) to provide "augmented human intelligence" services to users (Reynolds 2017). However, the use of individual techniques, for example deep learning neural networks or reinforcement learning, has found significant success across industry and applications (Rana and Oliveira 2015, Liu, Wang et al. 2017, Choy, Khalilzadeh et al. 2018).

Recently, there has been a proliferation of machine learning techniques and growing interest in their applications in finance, where they have been applied to sentiment analysis of news, trend analysis, portfolio optimization, risk modelling among many use cases supporting investment management. This chapter explores the potential of machine learning to enhance the investment process. A broad survey of the area was first conducted to determine the main programming languages, frameworks and use cases for machine learning from the perspective of the financial industry. Focus was then shifted to machine learning and its potential applications in quantitative investment – examining research that has applied machine learning to the investment process, analysing the technologies used, the functions of the applications, and evidence of potential to improve investment outcomes. The findings are relevant to both academics and practitioners with interest in investment management, and in particular quantitative investment, by providing a detailed discussion of the latest technologies, their potential uses, and probability of successful application.

The chapter is organized as follows. Section 2.2, provides an overview of the development of the area as a background for the discussion, this includes the emergence of machine learning, common algorithms and methodologies, and a review of the evolution and theory of quantitative investing. The research methods applied are discussed in Section 2.3. Section 2.4 provides a detailed description of the current state of the art in the application of machine learning to investment. Section 2.5 concludes with a discussion of the evidence presented.

2.2. Background

2.2.1. Machine Learning

Although variations of machine learning have long been around, the discipline has developed rapidly in recent years. Many factors have combined to derive this development. Increased computer power has made real time processing feasible for many complex tasks, increased connectivity has driven innovation and automation in the delivery of traditional tasks and services, the potential to extract useful information from the vast amounts of data generated via the internet (Big Data) has led to novel analytical methods. Alongside this, the development of easy to use programming languages, such as Python and R, and machine learning focused frameworks such as TensorFlow, has contributed to the wide investigation of machine learning applications in industry. It has already found commercial application across multiple industries from automated trading systems in the finance industry to the health sector where machine learning algorithms assist decision making in fertility treatments (Witten, Frank et al. 2016). The success of these applications is driving commercial research into further applications.

2.2.2. Common Machine Learning Approaches and Algorithms

Three main approaches to training machine learning algorithms are recognised; supervised learning, unsupervised learning and reinforcement learning. Supervised learning generates a function that maps inputs to outputs based on a set of training data. The algorithm infers a function linking each set of inputs with the expected, or labelled, output in the training set. Unsupervised learning finds hidden patterns in and draws inferences from unlabelled data. Unsupervised learning provides inputs to models, but does not specify an expected set of outcomes, the outcomes are unlabelled.

Reinforcement learning enables algorithms to learn by trial and error, based on feedback from past experiences. Like unsupervised learning, it does not require labelled data. A hybrid system, semi-supervised learning, combines supervised and unsupervised learning, using both labelled and unlabelled data to train models. This is useful where there is limited data or the process of labelling data could introduce biases.

The main research areas in supervised learning are regression and classification (specifying the category or class to which something belongs), this approach is often used in developing predictive models. Regression techniques predict continuous responses using algorithms such as linear regression, decision trees and Artificial Neural Networks (ANNs). Classification techniques predict discrete responses using algorithms such as logistic regression, Support Vector Machines (SVMs) or K-Nearest Neighbors (KNN). The main research area in unsupervised learning is clustering. Clustering refers to grouping objects together, such that objects that are put in the same group are more similar to each other than objects in other groups.

Artificial neural networks have become a key technology in the development of machine learning. They were first proposed over 75 years ago, inspired by the workings of the human brain (McCulloch and Pitts 1943). They are a collection of algorithms that replicate the process of a biological brain at the neuron level (Domingos 2012).

There are a number of different classes of artificial neural networks, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and recursive neural networks, among others. CNNs are ideal for things such as image classification and video processing because they're able to identify patterns by focusing on fragments of images. RNNs are better for dealing with things like speech or text analysis because they use time-series information, such as monthly stock price figures to predict next month's figure. GANs have garnered much interest in recent years since they were first introduced in 2014 (Goodfellow, Pouget-Abadie et al. 2014). GANs are comprised of two neural networks that compete against each other.

One neural network generates data similar to the training dataset, and the other tries to evaluate whether data is from the training dataset or generated by the generative network.

Aside from neural networks other well-known machine learning algorithms include SVMs, KNN and other. SVMs, used for classification and regression analysis, involve finding a hyperplane which minimizes the distance between a set of data points in an n-dimensional space. Bayesian networks are built from probability distributions and use probability laws for prediction and anomaly detection. KNN selects the most similar data points in the training data, this allows the algorithm to classify future data inputs in the same way. Some techniques are better suited to particular tasks than others. This research partly seeks to contribute to this area of knowledge. It is important to evaluate the effectiveness of certain algorithms, to assist in choosing appropriate algorithms for specific tasks in future applications and studies.

2.2.3. The Evolution of Quantitative Investing

Graham and Dodd's *Security Analysis*, published in 1934 following the Wall Street Crash of 1929 is the seminal work on fundamental investing and remains in publication today (Graham and Dodd 2008). It is one of the first books to distinguish investing from speculation, advocating the use of a systematic framework for analysing securities for stock selection.

A systematic approach to portfolio construction and risk analysis was presented in *Portfolio Selection* (Markowitz 1952). In this, Markowitz provides a mathematical definition of risk as the standard deviation of return. The approach focused on maximizing portfolio performance by optimizing the trade-off between risk and return. This was the foundation of modern portfolio theory, providing an analytical framework for the construction and analysis of investment portfolios (Kahn 2018) (Becker and Reinganum 2018).

A quantitative approach to market analysis gained popularity as advances in computing technology made the collection and analysis of large amounts of market data possible. This allowed the development and verification of market models on a scale not previously possible, contributing to significant advances in the understanding of financial markets, including the Capital Asset Pricing Model (CAPM) (Sharpe 1964, Mossin 1966, Lintner 1975, French 2003) and Efficient Market Hypothesis (EMH) (Malkiel and Fama 1970).

Fama and MacBeth (1973) used the Center for Research in Security Prices (CRSP) financial dataset (one of the first of its kind) to perform an empirical analysis of the CAPM. They showed that the CAPM provided a good quantitative approximation of the behaviour of security prices while setting a standard for empirical cross-sectional analysis of market data.

The empirical support for the EMH, enhanced by the success of market indices, such as the S&P 500, led to the dominant view, particularly in academia, that active investing was futile, as it was impossible to beat a passive investment. In comprehensive literature reviews, Kahn (2018) and Becker and Reinganum (2018) provide evidence that research and empirical evidence that challenged the CAPM and EMH was strongly discouraged. At the same time many examples of research that argued that although difficult, it is possible for active management to beat passive management, by exploiting market inefficiencies not covered by the CAPM and EMH. Strategies based on risk factor models, first explored by Rosenberg (1974) and Ross Ross (2013) in the 1970s, surged in popularity (Cochrane 2011) after the publication of the Fama-French three-factor model (Fama and French 1993).

From Markowitz portfolio optimization to CAPM, EMH and factor models more recently, quantitative investors have shown that they are willing to embrace new techniques and strategies. A key argument for applying machine learning techniques to financial problems is that machine learning methods capture non-linear relationships in the data (Lopez de Prado 2016). Non-linear methods are required to model data where outputs are not directly proportional to the inputs (Bianchi, Büchner et al. 2018) and many traditional analysis methods assume a linear relationship, or a non-linear model that can be simplified to a linear model. Typical examples of well-established non-linear machine learning methods include KNN, and ANN [20].

Machine learning has been applied with positive results across many areas of quantitative investing, including portfolio optimization (Lopez de Prado 2016, Heaton, Polson et al. 2017), factor investing (Nakagawa, Uchida et al. 2018), bond risk predictability (Bianchi, Büchner et al. 2018), derivative pricing, hedging and fitting (De Spiegeleer, Madan et al. 2018), and back-testing (Lopez de Prado and Lewis 2018). Section 2.4 contains a comprehensive summary of papers where machine learning techniques are applied to areas of quantitative finance.

2.3. Methodology

Initially, a broad search was conducted to identify the major themes related to machine learning. This search yielded information on the popular use cases and technologies. This information informed a second, more focused investigation of relevant material. Here, the aim was to draw connections between popular use cases in finance and current machine learning techniques.

As quality and scope of published research can vary widely, measures were taken to reduce the possibility of including unreliable information in the final dataset. Before inclusion in the concept matrix, each paper was assessed on quality. This was achieved by using a variety of quality indicators including; the citation count, the quality of an institute's research activities associated with the paper, bias created from funding sources, and the impact factor of the journal.

An appropriate search strategy was devised and carried out based on the main topics that were identified during the first investigation of the literature. The arXiv and SSRN databases were searched to ensure that the most up-to-date research papers were included. However, as these are not peer-reviewed papers, extra care was taken to ensure that the papers were from reputable authors, focusing on the quality of authors' previous publications. The topic phrases used in search were "portfolio management", "stock market forecasting", and "risk management". All of these topic phrases were used in conjunction with the key phrase "machine learning" in an attempt to return only relevant research papers. The purpose of searching by topic was to identify which technologies are widely and effectively used within each area. As the aim is to evaluate the current state of the art, there was a need to ensure that only recent papers were included. Thus, only papers that were submitted in 2015 or later were. From the initial search a total of 118 papers were collected. After an initial review of abstracts, papers that were not relevant to machine learning in finance (specifically investing) were removed. Any papers that were duplicates under more than one search topic were kept under the topic that appeared most relevant. Papers were then assessed in relation to their quality using the quality indicators mentioned above. This reduced the number of papers to 55.

2.4. Results

2.4.1. Popular Machine Learning Use Cases and Algorithms

A concept-centric matrix was utilised initially to identify which areas commonly use machine learning techniques. Recurring concepts and themes were noted and counted across a sample of 67 papers identified. An initial list of recurring themes was identified and analysed. Some themes, such as 'Geopolitics' were removed as they were deemed irrelevant due to the lack of research on the topic. A list of the most recurring themes with relevance to machine learning is presented in Table 2.4.1.

Table 2.4.1: Recurring themes from the literature review

Theme	**References**
Return Forecasting	21
Portfolio Construction	12
Ethics	8
Fraud Detection	8
Decision Making	8
Language Processing	7
Sentiment Analysis	7

The most common use-cases identified were return forecasting and portfolio construction. Quantitative methods were introduced to finance through the equity market and it is unsurprising that it should lead the way in incorporating the latest advances in its processes. A large number of the papers above also discussed risk modelling. This led us to take return forecasting, portfolio construction, and risk modelling as our three core topics. The most popular machine learning techniques identified in the papers researched are presented in Table 2.4.2, as well as a breakdown of the different acronyms used in the table.

Table 2.4.2: Popular techniques featured in machine learning and finance papers

	MLP	**SVM**	**LSTM**	**GRU**	**RNN**	**CNN**	**RF**	**GPR**	**LR**
Return Forecasting	7	5	4	2	-	1	2	-	-
Portfolio Construction	7	2	3	1	1	1	4	2	1
Risk Modelling	6	2	2	1	1	1	4	3	4

MLP Multilayer Perceptron
SVM Support Vector Machine
LSTM Long Short-Term Memory
GRU Gated Recurrent Unit
RNN Recurrent Neural Network (basic)
CNN Convolutional Neural Network
RF Random Forests/Decision Trees
GPR Gaussian Process Regression
LR Logistic Regression

Many techniques used in the papers only appear once, some twice. Since the purpose of this chapter is to identify the most popular machine learning techniques used in finance, specifically in the topics above, only techniques which appeared in at least three papers were included in Table 2.4.2. It was also decided to include RNN, although it is only mentioned explicitly in two papers, it appears implicitly more frequently as both LSTM and GRU are subsets of the technology.

Artificial neural networks are used in all three areas of finance studied, with a standard feedforward network (MLP) being the most common. Useful results are found from networks that range from small to very large networks (deep neural networks). There is also evidence of preferences for some techniques in particular areas. For example, Gaussian process regression is used in both portfolio construction and risk modelling but has not been applied to return forecasting.

2.4.2. Summary of Key Insights from Recent Papers

The paper selection included machine learning papers published in recent years as well as papers yet to be published by established authors from reputable institutions. These papers have been submitted for publication and are awaiting acceptance. The most recent studies in this field were included to help evaluate the cutting edge and state of the art of the use of machine learning for financial applications.

A. Portfolio Construction

Portfolio construction is the process of combining return forecasts and risk models to create an optimum portfolio given an investor's constraints. A variety of ANN methodologies are applied to the portfolio optimisation problem, often outperforming traditional optimisation techniques. Deep learning reappeared a number of times during this search in the context of portfolio construction. Deep learning refers to models that consist of multiple layers or stages of nonlinear information processing (for example, a neural network with many hidden layers) (Deng and Yu 2014). Both hierarchical clustering and reinforcement learning were used to improve portfolio diversification. Multiple papers discuss the method of applying Markov models to predict the performance of stocks. Markov models are a type of machine learning method that model variables that change randomly through time. The complicated nature of the global market makes using this type of model a viable option.

- The authors present a deep learning framework for portfolio design, applying their framework to the stocks in the IBB index, demonstrating that their portfolio weighted using deep learning outperformed the index (Heaton, Polson et al. 2017).

- The author outlines a reinforcement learning solution for a rational risk-averse investor seeking to maximize expected utility of final wealth, giving an example of a Q-learning agent exploiting an approximate arbitrage in a simulation (Ritter 2017).

- The authors of both papers make use of hierarchical clustering algorithms for constructing diversified portfolios. The portfolios are constructed using variations of risk parity (Lopez de Prado 2016) and equal risk contribution methods (Raffinot 2017) which take the hierarchical correlation structure of the assets into account. The portfolios constructed are shown to have superior diversification and out-of-sample risk adjusted performance.

- The authors make use of convex analysis techniques to devise an optimal portfolio coupled with a Hidden Markov Model (HMM) used to estimate growth rates in the market model, which achieves improved results over a simple model using geometric Brownian motions (Al-Aradi and Jaimungal 2019).

- The authors provide an overview of the financial applications of Gaussian processes and Bayesian optimisation, providing examples for forecasting the yield curve with Gaussian processes, and using Bayesian optimisation to build an online trend-following portfolio optimisation strategy (Gonzalvez, Lezmi et al. 2019).

- The authors compare the use of Feature Salient Hidden Markov Models (FSHMM) and HMM for constructing factor investing portfolios. The

FSHMM selects relevant factors for use from a pool of available factors, while the HMM uses the whole pool of factors. Both models outperformed benchmark portfolios, with the FSHMM portfolio showing better performance (Fons, Dawson et al. 2019)

- The authors use factors as inputs to deep neural network, SVM and random forest models for predicting stock returns. While their research shows the effectiveness of a deep learning model, more significantly they used Layer-wise Relevance Propagation (LRP) to determine individual factor contributions to the neural network's prediction (Nakagawa, Uchida et al. 2018).

- The authors create a non-linear multi-factor model using LSTM to estimate the non-linear function. As in the previous paper the authors make use of LRP to identify which factors contribute to the model. The performance of the LSTM model is compared to the neural network model used in (Nakagawa, Uchida et al. 2018) and gives superior returns (Nakagawa, Ito et al. 2019).

- The authors examine the use of three deep reinforcement learning algorithms, Deep Deterministic Policy Gradient (DDPG), Proximal Policy Optimization (PPO) and Policy Gradient (PG), in managing a portfolio of assets in the Chinese stock market. They determine that training conditions used in game playing and robot control are unsuitable for use with portfolio management, finding that DDPG and PPO gave unsatisfying performance in the training process. They propose the use of adversarial training methods and employ a revised PG algorithm which outperforms a Uniform Constant Rebalanced Portfolio (UCRP) benchmark (Liang, Jiang et al. 2018).

- The authors employ models constructed using Gaussian processes and Monte Carlo Markov Chains which learn optimal strategies from historical data, based on user-specified performance metrics (e.g. excess return to the market index, Sharpe ratio, etc.). This approach addresses the inverse problem of Stochastic Portfolio Theory – devising suitable investment strategies that meet the desired investment objective, when initially given a user-defined portfolio selection. The models outperform the benchmark in-sample and out-of-sample for absolute terms (returns) and also after adjusting for risk (Sharpe ratio) (Samo and Vervuurt 2016).

- The author provides an machine learning framework for estimating optimal portfolio weights. They apply this framework using three machine learning methods – Ridge and Lasso regression, and two newly introduced methods; Principal Component regression, Spike and Slab regression. All methods outperform the mean-variance, minimum-variance, and equal weight portfolios. (Kinn 2018).
- The authors propose a way to find the risk budgeting portfolio by using optimisation algorithms to find a solution to the logarithmic barrier problem. They use algorithms such as cyclical coordinate descent, alternating direction method of multipliers (ADMM), proximal operators, and Dykstra's algorithm (Richard and Roncalli 2019).
- The authors present a financial-model-free reinforcement learning framework as a solution to the portfolio management problem. The study tests the proposed framework with the following neural networks: CNN, a basic RNN, and LSTM (Jiang, Xu et al. 2017).

B. Return Forecasting

Return forecasting, predicting the investment return from an asset or asset class, is central to investment management and features highly in the literature. Many types of ANN are tested on their ability to forecast returns. Deep neural networks, CNNs, LSTMs are all applied to the problem of return forecasting. In one theme, the new machine learning technology is applied to improve forecasts made from traditional inputs, such as fundamental accounting data or technical indicators. A second approach uses machine learning to extract new inputs from alternative data, such as sentiment from news data. Finally, authors predict movement at market level rather than at the level of individual securities, for example using machine learning to identify states.

- The authors use a CNN strategy to analyse and detect price movement patterns in high-frequency limit order book data. Multilayer neural network methods and SVMs were also considered. However, they conclude the CNNs provide better performance for this task (Tsantekidis, Passalis et al. 2017).

- The authors implement several machine learning algorithms to predict future price movements using limit order book data. They employ two feature learning methods: Autoencoders, and Bag of Features. They compare three different classifiers: SVM, a Single Hidden Layer Feedforward Neural Network (SLFNN), and an MLP. They test the performance of the classifiers with an anchored walk forward analysis, to determine if the models can capture temporal information, as well as a hold-out per stock method, to determine if the models can learn features that can be applied to previously unseen stocks. The results from the MLP are better than the other classifiers. However, the use of the Autoencoder and Bag of Features in combination with the MLP lead to fewer correct predictions (Nousi, Tsantekidis et al. 2018).

- The authors introduce a novel Temporal Logistic Neural Bag-of-Features approach, that can be used to tackle the challenges that come with data of a high dimensionality, in this case high-frequency limit order book data (Passalis, Tefas et al. 2019).

- The authors train a deep neural network on reported fundamental data from publicly traded companies (revenue, operating income, debt etc.). The model forecasts future fundamental data based on a trailing 5-years window. A value investing factor strategy based on forecasted fundamental data outperforms a traditional value factor investing approach with a compounded annual return of 17.1% vs 14.4% for a standard factor model (Alberg and Lipton 2017).

- The authors create a simple buy-hold-sell strategy to predict direction of movement for 43 CME listed commodities and FX futures based on an ANN trained on a multitude of features for each instrument designed to capture co-movements and historical memory in the data. An average prediction accuracy of 42% is achieved across all instruments, with higher accuracies achieved for certain instruments (Dixon, Klabjan et al. 2017).

- The authors use a random forest model to predict the direction of stock prices based on price information and a number of momentum indicators (Relative Strength Index, Moving Average Convergence Divergence, Stochastic Oscillator, Williams %R, On Balance Volume, and Price Rate of Change). The algorithm is shown to outperform existing algorithms found in the literature (Khaidem, Saha et al. 2016).

- The authors provide a sentiment analysis dictionary which they use to predict stock movements in the pharmaceutical market sector. With this model they achieve an accuracy of 70.59%. (Shah, Isah et al. 2018)

- The authors present a methodology to define, identify, classify and forecast market states. They use a Triangulated Maximally Filtered Graph network to filter information, and simple logistic regression for predicting market states. They compare five models, with a Gaussian Mixture Model as their baseline. All five models outperform the baseline in terms of risk/return significance (Procacci and Aste 2018).

- The authors compare five ANN models for forecasting stock prices: a standard neural network using back propagation, a Radial Basis Function (RBF), a General Regression Neural Network (GRNN), SVM Regression (SVMR), and Least Squares SVM Regression (LS-SVMR). However, they compare the models on just three stocks: Bank of China, Vanke A, and Kweichou Moutai. The standard neural network using back propagation outperforms all of the other models across all three stocks, in terms of both Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). (Song, Zhou et al. 2018)

- The authors use 25 risk factors as inputs to machine learning stock returns prediction models. Results show that deep neural networks generally outperform shallow neural networks, and the best networks also outperform representative machine learning models (Abe and Nakayama 2018).

- The author employs ANNs to predict product demand for weather sensitive products in Walmart stores around the time of major weather events (Taghizadeh 2017).
- The authors implement a Gaussian Naïve Bayes Classifier for prediction based on sentiment analysis of Twitter data. The data used was obtained from Twitter and pertained to the 2014 FIFA world cup. Their framework obtained an accuracy and Area Under the curve of the Receiver Operating Characteristic (AUROC) of around 80% and an 8% marginal profit when tested (Le, Ferrara et al. 2015).

C. Risk

Three different themes are identified under the broad heading of risk. The first attempts to employ machine learning to improve traditional measures of risk used in the mean variance framework. The second theme looks for companies at risk of default or bankruptcy. Techniques such as natural language processing are used to identify words that indicate higher risk. The final theme uses machine learning to develop hedging strategies. Some authors look at identifying what selection of machine learning methods is best for risk modelling problems.

- The authors use k-means clustering to construct risk models by clustering stock returns normalized by standard deviation squared and adjusted by mean absolute deviation using a method proposed in (Kakushadze and Yu 2016). They demonstrate that this machine learning approach outperforms statistical risk models (Kakushadze and Yu 2017) in quantitative trading applications (Kakushadze and Yu 2019).

- The authors present a framework for hedging a portfolio of derivatives in the presence of market frictions such as transaction costs, market impact, liquidity constraints or risk limits (Buehler, Gonon et al. 2019).

- The authors show how Gaussian Process Regression can assist in pricing and hedging a Guaranteed Minimum Withdrawal Benefit (GMWB) Variable Annuity with stochastic volatility and stochastic interest rate (Goudenège, Molent et al. 2019).
- The authors show that machine learning can be as effective as other existing algorithms at solving difficult hedging problems in moderate dimension. They use techniques such as a modified LSTM neural network to calculate their hedging strategies (Fecamp, Mikael et al. 2019).
- The authors aim to explore the optimal model for business risk prediction. They attempt to do this using XGBoost, and by simultaneously examining feature selection methods and hyper-parameter optimization in the modelling procedure (Wang and Ni 2019).
- The authors try to predict daily stock volatility using news and price data. Their model, which utilizes a Bidirectional Long Short-Term Memory (BiLSTM) neural network and stacked LSTM's, outperforms the well-known Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model in all sectors analysed (financial, health care, etc.) (Sardelich and Manandhar 2018).
- The authors exploit a heterogeneous information network of 35,657 global firms to improve the predictive performance for firms likely to be added to a blacklist. Blacklists are used to keep track of entities that have unacceptable problems, such as financial or environmental issues. Blacklists help keep portfolios profitable and "green". Their model consists of a simple MLP with thirty hidden units (Hisano, Sornette et al. 2018).
- The authors estimate corporate credibility of Chinese companies using a CNN and natural language processing. They use Latent Dirichlet Allocation to summarise the text of news articles and use a CNN to extract the most important words from each topic. The CNN learns how news articles may reflect the credibility of a company though the wording of articles and word

occurrence. They verify their model works by building a negative rating system and showing a correlation between their model's results and the negative rating (Zhang, Luo et al. 2018).

- The authors compare different strategies for solving a variation of the multi-armed bandit problem. In their version of the problem, the learner can pull several arms simultaneously, or none at all. This could easily be applied to assist in investment decisions. Out of the strategies compared, Bayes-UCB-4P and TS-4P perform the best (Achab, Clémençon et al. 2018).

- The authors compare several machine learning algorithms: Logistic Regression, K-Dimensional Tree (K-D Tree), SVM, Decision Trees, AdaBoost, ANN, and Gaussian Processes (GP) for forecasting business failures (corporate bankruptcy). Models are compared on datasets of manufacturing companies in Korea and Poland. All of the models are compared on their performance when combined with different dimensionality reduction techniques. The techniques used are: Principal Component Analysis (PCA), Linear Discriminate Analysis (LDA), Isometric Feature Mapping (ISOMAP), and Kernel PCA. On the Korean dataset, all models perform similarly. K-D Tree, SVM, and GP perform best over all of the dimensionality reduction methods used. On the Polish dataset, the linear regression model performs the best. Although having a lower accuracy than some of the other models, it is the best performing method when compared over other results such as precision, recall, F1 score, and AUC (Area Under Curve) (Chow 2018).

2.5. Discussion

2.5.1. Strategy Development and Analysis

The results of the literature search demonstrate that there is a wide range of machine learning techniques being successfully applied to many areas in the development of quantitative investing strategies, outperforming traditional benchmarks, previously used techniques and algorithms in many cases. Algorithms that assume a linear relationship between data can result in reduced accuracy. (Lopez de Prado 2016) highlights this issue in terms of many of the econometric models employed by finance academics and investment managers. The author argues for the use of more advanced mathematical models and machine learning techniques such as unsupervised learning that are capable of modelling complex non-linear relationships in financial systems.

Taking factor investing as an example of this, (Harvey and Liu 2017) and (Harvey, Liu et al. 2016) make use of statistical algorithms to show that many factors discovered over the last number of years (particularly those found using empirical evidence) can be considered inaccurate or invalid. In the aptly named paper, *Taming the Factor Zoo*, a double selection LASSO machine learning method was used to analyse the contribution and usefulness of individual factors amongst the large number available today (Feng, Giglio et al. 2017). LASSO (Least Absolute Shrinkage and Selection Operator) is a regression analysis method capable of reducing the dimensionality of a large sample while selecting variables significant to the final result (Belloni, Chernozhukov et al. 2014). In (Abe and Nakayama 2018) the author uses twenty-five factors as model inputs, comparing the use of shallow and deep neural networks, as well as SVMs and random forests for predicting stock returns, finding the deep neural networks (more layers) superior to the other methods. Using a similar approach (Nakagawa, Uchida et al. 2018) uses factors as inputs to deep neural network, SVM and random forest models for predicting stock returns. While their research again showed the effectiveness of a deep learning model, more significantly they used layer-wise relevance propagation to determine individual factors contributions to the neural network's prediction. In these cases, not only has machine learning been used to develop investment strategies, but also to detect which input features were significant and which were not.

2.5.2. The Use of Alternative Data

The use of machine learning for the analysis and application of alternative data for example, sentiment analysis, supply chain data etc. has opened up opportunities for new investment strategies. As seen in Table 2.4.1, sentiment analysis was identified as a popular use case for machine learning. (Becker and Reinganum 2018) provides a thorough overview of the growth of big data and sentiment analysis research over the last 30 years, highlighting the use of techniques such as NLP, SVMs and ANNs for the analysis of news, conference calls, reports, and social media activity. They concluded that to date, sentiment information has provided short-term, easy to exploit insights but long-term persistent insights are hard to achieve (falling in line with EMH). (Kahn 2018) acknowledges the effectiveness of big data for the modern fundamental investor, as it can provide insights and improve decision making by widening their research capabilities. This sentiment is echoed in (Lopez de Prado 2016) where the author makes reference to the recently emerged term “quantamental" – describing a fundamentally leaning investor who manages their portfolio based on data-driven insights provided by machine learning algorithms. Examples of machine learning and alternative data being applied together in the results section mainly fall under return forecasting or risk modelling, where decisions may be made based on good or bad news (Shah, Isah et al. 2018), weather (Taghizadeh 2017), or social media sentiment (Le, Ferrara et al. 2015).

2.5.3. Choosing Machine Learning Algorithms

It is important to understand the relevant factors that contribute to the choice of machine learning algorithms, given the wide range available. These factors include accuracy, training time, linearity, number of parameters, the number of features and the structure of the data (Barga, Fontama et al. 2015). Some systems do not need a high level of accuracy. Estimates may be sufficient, for example, when calculating different route times for a journey. Model training times can also vary hugely between algorithms, making some algorithms more appealing than others when under time constraints. Many algorithms assume a linear relationship between input and output (linear regression, logistic regression, SVMs). This can result in reduced accuracy when dealing with non-linear problems. The number of parameters an algorithm has can indicate its flexibility, but also indicates that more time and effort may be required to find optimal values for training the model. The number of features can also be overwhelming for some algorithms. This is particularly a problem with textual data, where the number of words in the dictionary vastly outweighs the number of words in say, a paragraph being used for sentiment analysis. It's important to consider the structure of the data and the specific problem, as some algorithms are better suited for certain problems and data structures (Harrington 2012).

2.5.4. Backtesting and Strategy Verification

While machine learning techniques can provide superior performance, financial data is notorious for having a low signal-to-noise ratio, which can lead to the detection of false patterns and results. Backtesting protocols have been proposed to tackle this (Arnott, Harvey et al. 2019). Machine learning solutions have also been applied to this problem. In (Lopez de Prado and Lewis 2018) the authors present an unsupervised learning strategy which makes use of a modified k-means clustering algorithm to extract the number of uncorrelated trials from a series of backtests, which can be used in estimating the probability of false positives and estimating the expected value of the maximum Sharpe ratio. While in (Wiecki, Campbell et al. 2016) the authors use a machine learning strategy for backtesting and the evaluation of automated trading strategies which is trained on a number of performance and risk metrics, demonstrating that this strategy outperforms standard metrics such as Sharpe ratio out-of-sample.

The development of new backtesting strategies and protocols is welcome and necessary, especially taking into account recent "black box" criticisms by leading deep learning researchers regarding a lack of testing and reproducibility in the field of machine learning. In their acceptance speech after winning the "test-of-time" award at NIPS, the leading AI conference, the authors of (Recht and Rahimi 2017) compared much of recent machine learning research to "alchemy", highlighting a situation where algorithms were being created and trained using trial and error methods, with the researchers unable to explain the fundamental operation. They later published a paper highlighting instances of this (Sculley, Snoek et al. 2018).

2.6. Conclusion

As the previous section discusses, machine learning offers an opportunity for more complex financial analysis than was previously possible. The literature shows that quantitative investors have embraced new tools and techniques as they have emerged (Becker and Reinganum 2018, Kahn 2018).

There is a growing body of literature applying machine learning techniques to investment problems. Varieties of machine learning methods have been applied to areas of quantitative finance– the most popular methods are MLPs, followed by SVMs, and LSTM. Machine learning has been applied to problems in areas such as return forecasting, portfolio construction, and risk modelling. These machine learning methods utilize traditional financial data, as well as making use of new types of alternative data. Big data is providing new datasets that need to be analysed and machine learning techniques are capable of modelling complex (non-linear) relationships and analysing new data.

Lopez de Prado (2016) notes the recent trend of traditional hedge funds hiring an increasing proportion of STEM graduates for portfolio construction positions, as they possess the required mathematical skillset for performing complex analysis and computer modelling. An understanding of machine learning, as well as the languages (Python, R, etc.) and frameworks (e.g. TensorFlow) needed to construct complex models could certainly be considered advantageous for any quantitative investor looking for an edge.

Chapter 3. Regime Switching Models

3.1. Introduction

Regimes are periods of time with unique characteristic financial variables, such as mean returns, correlations and volatilities. A change in regimes implies a change in the characteristic behaviour of the financial market which may continue for some time if the regime is persistent.

Regimes are an easy concept to grasp intuitively, a distinction can be made between "Bull" and "Bear" markets, periods of high and low volatility returns, as well as periods of change in policy or regulation (Ang and Timmermann 2012). Regime switching Hidden Markov Models (HMM) have been shown to capture regimes over various time horizons and have been successfully implemented across a range of asset allocation strategies.

Regime switching models first gained popularity in quantitative finance when they were introduced in Hamilton's (1989) seminal work which used a HMM to identify expansions and recessions in the business cycle. A hidden Markov model is an unsupervised machine learning technique is used to infer regimes. Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labelled responses.

In this chapter, the properties of regime switching HMMs are discussed, as well as their use in asset allocation strategies. The literature shows that regime switching HMMs are a useful and powerful tool for analysing financial data. Model selection criteria and the implementation of HMMs are also discussed.

.

3.2. Capturing Stylized Facts using HMMs

Financial returns have been shown to exhibit distributional properties such as skewness and leptokurtosis, making the normal distribution unsuitable for their description (Cont 2001). HMMs using both normal and other distributions, e.g. student-t distributions, have been shown to more effectively model the stylized facts of returns by identifying a number of market regimes (Hamilton and Susmel 1994, Rydén, Teräsvirta et al. 1998, Bulla and Bulla 2006, Ang and Timmermann 2012).

The main argument for the use of market regimes and regime-switching in asset-allocation strategies is that all-weather portfolios and static asset-allocation strategies cannot account for the presence of all possible market regimes (De Prado 2019).

3.3. Regime Switching in Asset Allocation

The ability to infer regime changes has been shown to be profitable across a range of asset allocation strategies. Ang and Bekaert (2002) concluded that a high-volatility, high-correlation regime is present in a regime-switching model during a bear market. Following on from this finding, they demonstrated that a regime-switching based dynamic asset-allocation strategy outperforms a static strategy by switching to cash during high-volatility regimes (Ang and Bekaert 2004).

Guidolin and Timmermann (2008) consider an international asset-allocation strategy where a regime-switching model is used to capture periods of distinct skewness and kurtosis in global equity returns. They identify home bias among US-based investors, justifying this bias during certain regimes where the US stocks have high skewness and low kurtosis compared to the global market portfolio.

Bulla, Mergner et al. (2011) demonstrated that a regime-based asset allocation strategy under realistic assumptions could outperform a buy and hold strategy after taking transaction costs into account.

More recent papers have demonstrated similar positive results using dynamic asset allocation strategies, using a HMM with time-varying parameters (Nystrup, Hansen et al. 2015, Nystrup, Hansen et al. 2017)

While the strategies above focus on modelling returns, regime switching asset allocation strategies have also been developed using other sources of data. Ammann and Verhofen (2006) examined a multivariate regime-switching model estimated using the Carhart (1997) four-factor model, finding two distinct regimes, a high-variance regime where value stocks performed well, and a low variance regime where momentum stocks were the highest performer. Kritzman, Page et al. (2012) defined a regime-switching model based on a number of economic variables; market turbulence, inflation, and economic growth. A dynamic asset-allocation strategy based on this regime-switching model outperformed static asset allocation.

3.4. Model Selection

It has been observed in literature reviews that a large proportion of papers relating to regime-switching models focus on two regime models, generally thought to represent "bull" and "bear" markets, or "good" and "bad" regimes (Ang and Timmermann 2012). Guidolin (2011) observed that approximately half of the literature refers to two regime models. He also noted that the literature was roughly split in half between papers taking a statistical approach of "letting the data speak", and papers where regime-switching models were used as a tool to give a plausible explanation based on underlying economic theory and reasoning. Papers based on economic reasoning were more likely to restrict the scope of the research to a two-regime model than papers taking a statistical approach, where selection of a model appropriate for the data was prioritised.

Guidolin and Timmermann (2007) defined a regime-switching model characterized by four states; crash, slow growth, bull and recovery. The model was estimated using stock and bond returns. Each state had a distinct optimal asset allocation, and an out-of-sample test confirmed the effectiveness of the model. Maheu, McCurdy et al. (2012) proposed a four-regime model where the regimes could be characterised as bull, bull correction, bear, and bear rally, arguing that this provided a richer characterization of market cycles. Their model was shown to outperform alternative approaches including a two-regime model, showing that the four-regime model was less prone to erratic switching, giving more persistent regimes.

Regime persistence, and the number of state changes has implications for the practicality of regime-switching models, as potential excess returns can be reduced by rebalancing costs if switching behaviour is too noisy or frequent (Bauer, Haerden et al. 2004, Hess 2006). Filtering procedures can be used to smooth the sequence of predicted regimes, decreasing transaction costs and increasing regime persistence. Bulla, Mergner et al. (2011) applied a median filter to the sequence of predicted regimes across a number of major indices, significantly impacting transaction costs as the filter reduced the number of regime changes by 50-65%.

It has been acknowledged that increasing the number of regimes can lead to overfitting and a quadratic increase in the number of parameters of the regime-switching model which need to be estimated. Penalised likelihood criteria such as the Bayesian information criteria (BIC) or Akaike information criteria (AIC) can be used to estimate the optimum number of regimes with BIC being considered more reliable (Guidolin 2011, Gatumel and Ielpo 2014, Nystrup, Madsen et al. 2015).

Gatumel and Ielpo (2014) modelled the dynamics of stocks, bonds, commodities and currencies using regime-switching models and rejected the hypothesis that two regimes were enough to model the behaviour of asset returns. The empirical results of their testing suggested that to capture the distributional characteristics of the various assets required between two and five regimes.

Nystrup, Madsen et al. (2015) compared standard HMMs, semi-Hidden Markov Models (HSMM), and a new method, continuous time HMMs (CTHMMs) for describing the stylized facts of daily returns in the S&P 500 index. HSMM were first introduced in Bulla and Bulla (2006) and were shown to describe stylized facts well, outperforming HMMs in some cases. CTHMMs tackle the issue of parameter growth in HMMs, with parameters increasing linearly rather than quadratically as the number of regimes increases, which decreases the likelihood of overfitting. BIC and AIC indicated that four-regime CTHMMs and HSMMs were best suited to the data. Testing demonstrated that CTHMMs better captured the stylised facts of asset returns, with a four-regime model outperforming two or three regime models across all three classes of model.

3.5. Hidden Markov Models

As shown in the discussion above, regime switching HMMs are a useful tool for analysing and capturing the properties of financial data. In this section the operation of the HMM is defined and explained.

The main characteristic of a hidden Markov model is a probability distribution of the observation $X_t,\ t = 1, \ldots, T$ which is dependent on the states S_t of an unobserved first-order Markov chain.

A sequence of discrete random variables $\{S_t: t \in \mathbb{N}\}$ is said to be a first-order Markov chain if, for all of $t \in \mathbb{N}$, it satisfies the Markov property

$$\mathrm{P}(S_{t+1} | S_t, \ldots, S_1) = \mathrm{P}(S_{t+1} | S_t) \tag{1}$$

A transition probability matrix (TPM) governs the switching behaviour of the model between states. In a two-state model for example, the TPM would be of the form

$$\Pi = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \tag{2}$$

with $p_{ij}, i, j \in \{1,2\}$ denoting the probability of being in state j at time $t + 1$, given a sojourn in state i at time t. The observation X_t has a distribution at time t specified by $\mathrm{P}(X_t = x_t | S_t = s_t)$, the conditional or component distributions of the model. A two-state Gaussian component distribution would give

$$x_t = \mu_{s_t} + \epsilon_{s_t}, \epsilon_{s_t} \sim N\left(0, \sigma_{s_t}^2\right) \tag{3}$$

Where $\mu_{s_t} \in \{\mu_1, \mu_2\}$ and $\sigma_{s_t}^2 \in \{\sigma_1, \sigma_2\}$. To estimate the parameters of the HMM an expectation-maximization (EM) method such as the Baum-Welch algorithm is commonly used (Baum, Petrie et al. 1970).

Assuming that successive observations are independent, the likelihood function is given by:

$$L(\theta) = \boldsymbol{\pi P}(x_1)\Pi \boldsymbol{P}(x_2)\Pi \dots \boldsymbol{P}(x_{T-1})\Pi \boldsymbol{P}(x_T)\mathbf{1} \quad (4)$$

where $\boldsymbol{P}(x_t)$ is a diagonal matrix containing the state-dependent conditional distributions as entries and $\boldsymbol{\pi}$ denotes the initial distribution of the Markov chain. After parameter estimation of the HMM, the hidden states can be inferred. A common technique for determining the most likely sequence of states is the Viterbi algorithm (Viterbi 1967). The algorithm calculates the most probable sequence of states using:

$$\{\hat{\mathrm{s}}_1, \dots, \hat{\mathrm{s}}_\mathrm{T}\} = \underset{j_1,\dots,j_T}{\operatorname{argmin}} P(S_1 = j_1, \dots,, S_T = j_T | X_1^T = x_1^T) \quad (5)$$

Chapter 4. A Flow-based Regime-Switching Model: Data & Methodology

4.1. Introduction

As discussed in the previous chapter, regime-switching models are a useful machine learning tool for modelling and analysing financial data. In this chapter, regime-switching models are applied to analyse international portfolio flow data.

International portfolio flows describe the actions of informed investors, shareholders, and fund managers who add or remove cash from funds and buy or sell individual securities with fund deposits. Flows have been shown to be stable and persistent in nature. They have also been shown to influence equity returns. Based on the study of previous literature relating to regime-switching models, and portfolio flows, several methods were used examine the properties of flows using regime-switching models, and to examine the behaviour of regime-switching models estimated using flow data.

To confirm the persistent nature of flows, the transition probability matrices of regime-switching models estimated using flows were examined and compared with models estimated using returns. Examining the switching behaviour of models over time can be used to identify interesting features in the data such as structural breaks.

To examine the relationship between returns and flows, returns data corresponding to time periods when certain regimes were present was used to characterise each regime based on returns as well as flows. Characteristic returns that are unique for each regime and from the overall dataset would indicate that there is a relationship between returns and flows. If it is possible to identify characteristic returns using a flow-based regime-switching model, this information can be used as part of a regime-based asset allocation strategy. To test this, a walk-forward out-of-sample test is proposed for a portfolio of global equity indices, where long or short positions are taken based on a signal from a regime-switching model.

4.2. International Portfolio Flows

International portfolio flows have been shown to affect equity prices in developed and emerging markets. Quarterly cross-border equity inflows and outflows were found to cause international prices to rise and fall respectively (Tesar and Werner 1994, Tesar and Werner 1995, Brennan and Cao 1997). Froot, O'Connell et al. (2001) found that flows appear stationary and more persistent than returns, while also finding that flows have an influence on returns. Also using daily data, Griffin, Nardari et al. (2004) found that equity inflows into a country increase with that country's stock market returns. For small countries, equity flows increase when U.S. and world market returns increase, irrespective of local market performance. Froot and Ramadorai (2008) show that weekly cross-border equity flows forecast emerging market equity returns as well as suggesting that predictions of domestic equity returns using flows are due to information rather than price pressure. Froot and Teo (2008) demonstrate that flows capture factor investing behaviour among institutional investors across three areas: size, value/growth and sector. They also show that factor flows influence and forecast returns. Froot, Bhargava et al. (2014) aggregate flows across various asset classes and demonstrate that these aggregated flows can be used to measure market sentiment, improve the timing of asset class specific risk, and inform asset allocation strategies.

4.3. Flow Data

The portfolio flows used in this research are provided by State Street Global Markets (SSGM). The flows are constructed using proprietary capital allocation data from thousands of institutional investor portfolios under custody and administration by State Street Corporation (SSC), one of the largest custodian banks in the world with US$31.62 trillion assets under custody and US$2.51 trillion assets under management (State Street Corporation 10-K, 2019).

The flow data used is part of a series of Equity Flow Indicators (EFI) developed by SSGM, published at daily frequency since March 1998. Flows are available at global and regional levels, as well as for countries and industry sectors - regional and country flows were used in this research. Each EFI is comprised of two components, a benchmark flow, and an active flow. These two components are summed together to give the total flow. The benchmark flow is constructed empirically based on capital from fund deposits being allocated to securities according to the weights of existing portfolio positions. The active flow is taken as the difference between the total observed flow and the benchmark flow. This can be considered as representing actions by fund managers that deviate from the benchmark, for example buying or selling individual securities with fund deposits. Therefore, active flows can be considered as describing portfolio rebalancing, while benchmark flows describe portfolio resizing. Total flows, the sum of the benchmark and active flows is used in this research. Froot, O'Connell et al. (2001) contains a detailed analysis and description of the SSC flow dataset.

To measure flows, individual security flows are combined across portfolios and measured relative to AUM or market capitalisation. Aggregated flows are created by summing the individual security flows across the appropriate country, region or sector. The EFI is the ten-day exponentially weighted moving average of these flows.

The flows used in this research are AUM-weighted, market cap normalized and calculated using the following method.

$$flow_{s,t} = \sum_{funds} flow_{f,s,t} \tag{6}$$

$$Flow_{M,t}^{Mcap,AUMWtd} = \frac{1}{MCap_{M,t}} \sum_{s \in M} flow_{s,t} = \frac{flow_{M,t}}{MCap_{M,t}} \tag{7}$$

Where the AUM-weighted flow for security s and fund f is calculated at time t and normalized by the market cap $MCap$ of a given market m.

Cross-border EFI data corresponding to the eight countries and regions shown in Table 4.3.1 was obtained from SSGM for the time period January 2001 to December 2018 on a daily frequency. Weekly flows were calculated as the net daily flows for that week. Descriptive statistics for daily and weekly flows are shown in Table 4.3.2 and Table 4.3.3.

4.4. Price Data

MSCI Total Return index data was obtained from Thomson Reuters Datastream for the corresponding country and regional indices shown in Table 4.3.1 for the time period January 2012 to December 2018. Returns were calculated for the indices using:

$$r_t = \log P_t - \log P_{t-1} \tag{8}$$

Descriptive statistics for daily and weekly returns are shown in Table 4.3.2 and Table 4.3.3. The daily 1-month USD LIBOR was obtained from Thomson Reuters Datastream for the time period January 2012 to December 2018. This was used as the risk-free rate when calculating the Sharpe ratio for the portfolios in Section 5.2

Table 4.3.1: Investment Universe

Investment Universe	
USA	Pacific ex Japan
UK	EM Asia
Japan	EM EMEA
Europe Ex UK	EM Latin America

Table 4.3.2: Descriptive statistics of daily returns and daily flow data for the period January 2012 to December 2018, displaying the mean (μ) and standard deviation (σ).

Descriptive statistics of daily returns and flows (Daily Data, 2012-2018)

	Daily Returns (%)		Daily Flows	
	μ_r	σ_r	μ_f	σ_f
Europe ex UK	0.037	0.945	0.015	0.067
EM Latin America	0.028	0.934	0.018	0.141
EM EMEA	0.032	0.848	0.029	0.068
Pacific ex Japan	0.037	0.698	0.005	0.052
EM Asia	0.030	0.776	0.017	0.059
USA	0.049	0.798	-0.048	0.078
UK	0.028	0.836	0.012	0.108
Japan	0.054	1.221	-0.002	0.059

Table 4.3.3: Descriptive statistics of weekly returns and weekly flow data for the period January 2012 to December 2018

Descriptive statistics of weekly returns and flows (Weekly Data, 2012-2018)

	Weekly Returns (%)		Weekly Flows	
	μ_r	σ_r	μ_f	σ_f
Europe ex UK	0.162	2.046	0.077	0.314
EM Latin America	0.109	2.053	0.091	0.673
EM EMEA	0.132	1.914	0.145	0.310
Pacific ex Japan	0.157	1.602	0.024	0.245
EM Asia	0.126	1.904	0.088	0.277
USA	0.220	1.664	-0.241	0.375
UK	0.116	1.803	0.059	0.516
Japan	0.216	2.810	-0.009	0.268

4.5. Training & Characterising Regimes using Flow data

4.5.1. Persistence of Flows vs Log Returns

Estimating the parameters of a HMM using an EM algorithm gives Π, the transition probability matrix (TPM). A TPM with a strong diagonal ($\Pi_{\mathrm{ii}} \gg 0.5$) indicates that the regimes will be highly persistent i.e. that the regime inferred by the model at t will most likely continue at $t + 1$.

By examining the TPM diagonals of regime-switching models estimated using flow data and comparing them to the diagonals for models trained using log returns over the same time periods and frequencies we can compare the persistence of flow-based and returns-based regimes. Persistent regimes have many practical benefits in terms of reducing transaction costs and reducing the need for probability smoothing and filtering.

For example, training a 4-regime HMM over the period January 2012 to December 2018 using daily cross-border equity flow data, gives a TPM with a very strong diagonal ($\Pi_{\mathrm{f}_{ii}} > 0.9$).

$$\Pi_{\mathrm{f}} = \begin{pmatrix} 0.952 & 0.000 & 0.027 & 0.021 \\ 0.002 & 0.969 & 0.018 & 0.012 \\ 0.017 & 0.041 & 0.919 & 0.024 \\ 0.013 & 0.023 & 0.018 & 0.946 \end{pmatrix}$$

When a HMM is trained over the same period January 2012 to December 2018 using the daily log returns of the MSCI Total Return indices for the eight regions, the TPM has a much weaker diagonal. While $\Pi_{\mathrm{r}_{ii}} < 0.5$ for three of the regimes, indicating that they are persistent, the TPM of the HMM trained on flows is much stronger.

$$\Pi_{\mathrm{r}_{\mathrm{ii}}} = \begin{pmatrix} 0.440 & 0.075 & 0.409 & 0.077 \\ 0.088 & 0.665 & 0.232 & 0.015 \\ 0.182 & 0.096 & 0.0712 & 0.010 \\ 0.124 & 0.022 & 0.012 & 0.842 \end{pmatrix}$$

TPM diagonal values for regime switching models with k number of regimes were estimated using flow data and log returns data on both a daily (Table 5.1.1) and weekly (Table 5.1.3) frequency.

4.5.2. Identifying & Characterising Regimes present in Flow Data

The Viterbi algorithm was used to infer the most likely sequence of regimes present in the flow data. This allowed for the switching behaviour of the models to be analysed over time, as well as allowing for the regimes to be characterised based on returns.

Examining the switching behaviour of the model over time allows us to visually inspect the regimes assigned to the data and identify characteristics of the model such as structural breaks, where there is a long-term change or permanent change i.e. a regime would no longer appear in the model. Initially, the most likely sequence of regimes was estimated across the whole dataset, from January 2001 to December 2018, and there appeared to be a structural break across several models around the end of 2011. For the purposes of a regime-based asset allocation strategy, it was decided to only train the model on data from January 2012 onwards to avoid scenarios where for example, a 4- or 5-regime model would only really have two regimes present after 2012 to capture the behaviour of the data (discussed further in Section 5.1.2).

While portfolio flows describe the behaviour of investors, it is their relationship with returns that is of most interest. Therefore, based on the periods when a certain regime was present, the corresponding returns data for each instrument during those periods was used to characterise that regime based on mean returns and standard deviation (Section 5.1.3). The presence of unique returns characteristics for each regime would allow for an investor to capture excess returns by adjusting their portfolio for the presence of that regime.

4.5.3. Model Selection

Bayesian information criterion (BIC) and Akaike information criterion (AIC) are both model selection criterion based on using the likelihood function to score the model for overfitting based on the number of parameters used in the model.

Both BIC and AIC penalise the model for increasing the number of parameters in the model, as this tends to increase overfitting. The penalty term is larger with the BIC than with the AIC.

The BIC is defined as:

$$BIC = -2\log L + p\log T \tag{9}$$

The AIC is defined as:

$$AIC = -2\log L + 2p \tag{10}$$

Where L is the log likelihood for the model, T is the number of observations and p is the number of parameters.

To test the appropriate number of regimes for use in the regime detection model the BIC and AIC were calculated for hidden Markov models as the number of possible regimes was increased. The model with the lowest score should be selected.

4.6. A Regime-Based Asset Allocation Strategy

A HMM was used to infer regimes based on cross-border equity flow data for the eight regions in Table 4.3.1. To characterise the regimes based on price movement, mean returns were calculated for each regime using MSCI Total Return Index data corresponding to the eight regions. To examine the ability of flow-based regimes to capture returns, a portfolio was constructed with the eight MSCI indices as assets. A long or short position was taken in each asset based on their mean returns for the current regime according to the HMM. If flows, and through them, regimes are persistent, the positions will capture consistent returns.

To initially estimate the parameters of the HMM and to characterise the inferred regimes based on their mean returns, a training period of January 2012 to December 2016 was chosen (discussed in Section 4.5.2). The Viterbi algorithm was used to infer the most likely sequence of regimes up to the current regime.

Beginning with the regimes characterised in the 2012-2016 training period, an out-of-sample test was implemented from January 2017 to December 2018 using the following method which stepped through the data iteratively to exclude future information.

Using flow data from January 2012 up until time t, calculate the most likely sequence of regimes up until the current regime k.

The mean return is calculated for every instrument in the portfolio across the time periods spent in the current regime:

$$\mu_{t,k}^{i} = \frac{\sum_{t,k=1}^{T_k} r_{t,k}^{i}}{T_k} \tag{11}$$

Where $\mu_{t,k}^{i}$ is the mean return of instrument i at time t in regime k, T_k is the time spent in regime k and $r_{t,k}^{i}$ is the return of instrument i at time t in regime k.

Based on whether the mean return $\mu_{t,k}^{i}$ is positive or negative for the current regime k at time t, a simple long/short trading signal M_t^i is assigned to each instrument:

$$M_t^i = sign\left(\mu_{t,k}^{i}\right) \tag{12}$$

Each instrument in the portfolio is assigned an equal weight. To try and simulate a more realistic trading scenario the signal for time t is used to determine each weight at time $t + \tau$, where τ indicates a time shift. The weights are calculated as follows:

$$w_{t+\tau}^{i} = \frac{1}{N} M_t^i \tag{13}$$

Where $w^i_{t+\tau}$ is the weight of instrument i held in the portfolio at time $t+\tau$, N is the number of instruments in the asset class and M^i_t is the regime-based signal from the previous time period.

The return of the portfolio $r_{t+\tau}$ is calculated as:

$$r_{t+\tau} = \sum_{i=1}^{N} (r^i_{t+\tau}\, w^i_{t+\tau}) \tag{14}$$

Where $r^i_{t+\tau}$ is the return of instrument i at time $t+\tau$.

Chapter 5. A Flow-based Regime-Switching Model: Results

The results are comprised of two sections. The first section examines the characteristics of regime-switching models estimated using flow data. The persistence of flow-based regimes is compared to regimes estimated using returns. The switching behaviour of flow-based regimes is examined over the period 2001-2018, where the existence of persistent regime shifts and structural breaks are noted. Descriptive statistics are calculated for individual regimes allowing them to be compared in terms of both flows and returns. Model selection criteria and the optimal number of regimes are also discussed.

The second section focuses on the performance of a regime-based asset allocation strategy, showing the ability of flow-based regimes to capture excess returns. The results of walk-forward out-of-sample tests are given for models estimated using daily and weekly flow data. The optimal number of regimes in a model is compared in a series of robustness tests.

5.1. Training & Characterising Regimes using Flow Data

5.1.1. Persistence of Flows vs Returns

As discussed in section 4.5.1, the diagonal of the transition probability matrix (Π_{ii}) for a regime-switching model indicates the persistence of the regime i.e. the probability of remaining in the same regime at $t+1$ after being in that regime at t.

TPM diagonal values for k-regime models were estimated using flow data and log returns data on both a daily (Table 5.1.1) and weekly (Table 5.1.3) frequency.

Table 5.1.1 and Table 5.1.2 show the TPM diagonals for various regime-switching models estimated using daily flow data or daily log returns. The models estimated using flow data all showed $\Pi_{ii} > 0.9$, indicating that the flow-based regimes are highly persistent. For the models estimated using log returns, it was observed that as the number of regimes increases, the instances where Π_{ii} <0.5 also increases. While an increased number of regimes provide a better fit for the data in terms of capturing outliers and heavy-tailed distributions, the transient nature of these regimes are not practical in terms of timing and increased transaction costs. Implementing these models would most likely require additional filtering and probability smoothing which could negate the ability to capture a sudden change in regimes. The persistence of flow-based regimes reduces the need for additional filtering and smoothing, potentially reducing transaction costs.

Table 5.1.3 and Table 5.1.4 show the TPM diagonals for regime-switching models estimated using weekly data. In general, the values of Π_{ii} are lower for the models estimated on a weekly frequency compared to those estimated on a daily frequency. Intuitively, this makes sense as a change in regimes is more likely to occur between time periods on a longer time horizon. However, the observation that flow-based regimes are more persistent than regimes estimated using returns still holds true on a weekly frequency.

A. Daily Frequency

Table 5.1.1: Transition probability matrix diagonals (Π_{ii}) for regime switching models with k number of regimes. Model parameters estimated using daily flow data for the period January 2012 to December 2018.

Transition Probability Matrix Diagonal (Π_{ii})
Daily Flow Data Models

k-Regime Model	**Regime (i)**				
	1	**2**	**3**	**4**	**5**
1	1				
2	0.970	0.974			
3	0.966	0.975	0.951		
4	0.952	0.946	0.919	0.969	
5	0.956	0.972	0.922	0.926	0.940

Table 5.1.2: Transition probability matrix diagonals (Π_{ii}) for regime switching models with k number of regimes. Model parameters estimated daily log returns for the period January 2012 to December 2018.

Transition Probability Matrix Diagonal (Π_{ii})
Daily Log Returns Models

k-Regime Model	Regime (i)				
	1	2	3	4	5
1	1				
2	0.787	0.918			
3	0.373	0.696	0.815		
4	0.307	0.227	0.699	0.815	
5	0.248	0.812	0.231	0.554	0.199

B. *Weekly Frequency*

Table 5.1.3: Transition probability matrix diagonals (Π_{ii}) for regime switching models with k number of regimes. Model parameters estimated using weekly flow data for the period January 2012 to December 2018.

Transition Probability Matrix Diagonal (Π_{ii})
Weekly Flow Data Models

k-Regime Model	Regime (i)				
	1	2	3	4	5
1	1				
2	0.964	0.935			
3	0.968	0.857	0.896		
4	0.762	0.956	0.773	0.876	
5	0.692	0.840	0.931	0.797	0.750

Table 5.1.4: Transition probability matrix diagonals (Π_{ii}) for regime switching models with k number of regimes. Model parameters estimated using weekly log returns for the period January 2012 to December 2018.

Transition Probability Matrix Diagonal (Π_{ii})
Weekly Log Returns Models

k-Regime Model	**Regime (i)**				
	1	**2**	**3**	**4**	**5**
1	1				
2	0.599	0.864			
3	0.585	0.329	0.803		
4	0.301	0.652	0.284	0.322	
5	0.241	0.108	0.312	0.527	0.548

5.1.2. Model Switching Behaviour & Structural Breaks

Figures 5.1.1-4 below and Figures A.1.1-6 in Appendix 1, show the posterior probabilities over time - the probability that a regime was present at time t - for models estimated using daily and weekly flow data. During initial analysis of model switching behaviour over the period January 2001 to December 2018, a structural break appears in the flow data around the end of 2011. After this point certain regimes no longer appear to be present and one regime appears to dominate the switching behaviour.

In the case of a 2-regime model, the model remains in Regime 1 for almost the entire period after 2011, rarely switching back to Regime 2, which had been the dominant regime from 2008 until 2011 (Figure 5.1.1 & Figure 5.1.2).

When a 3-regime model is estimated over the same period the model again remains in Regime 1 for a long period from late 2011 to 2016, with Regime 2 being the dominant period from 2008 to 2011 (Figure A.1.1 & Figure A.1.2, Appendix 1).

The 4-regime and 5-regime models also appear to exhibit similar behaviour, remaining in Regime 1 for a long period from late 2011 to 2016. However, the period from 2008 to 2011 is further decomposed into Regimes 2 and 3 in the 4-regime model (Figure

A.1.3 & Figure A.1.4, Appendix 1), and Regimes 2, 3 and 4 in the 5-regime model (Figure A.1.5 & Figure A.1.6, Appendix 1).

A possible economic explanation that could be assigned to the persistent presence of certain regimes during the period 2008 to 2011, is the Global Financial Crisis. The period after this from 2011 to 2016 where there is an apparent structural break, could possibly be related to factors such as monetary policy and quantitative easing after the Financial Crisis. These, or other possible explanations for regime switching behaviour were not verified or explored further but could provide an interesting topic for further research.

Figure 5.1.1: Regime posterior probabilities over time for a 2-regime model estimated using daily data.

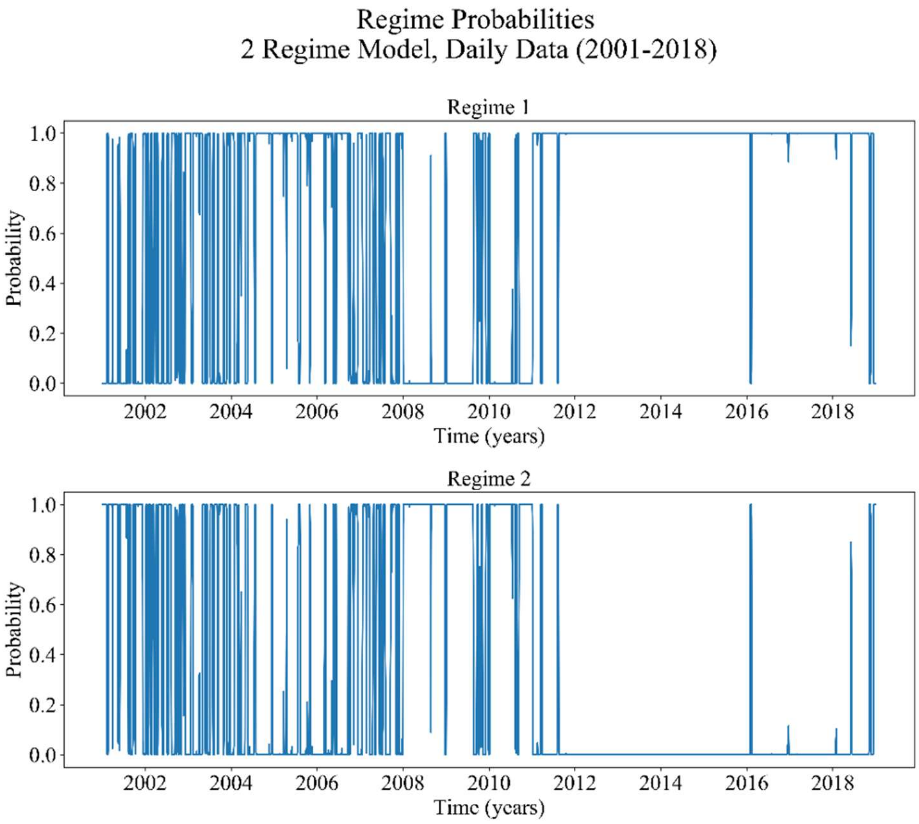
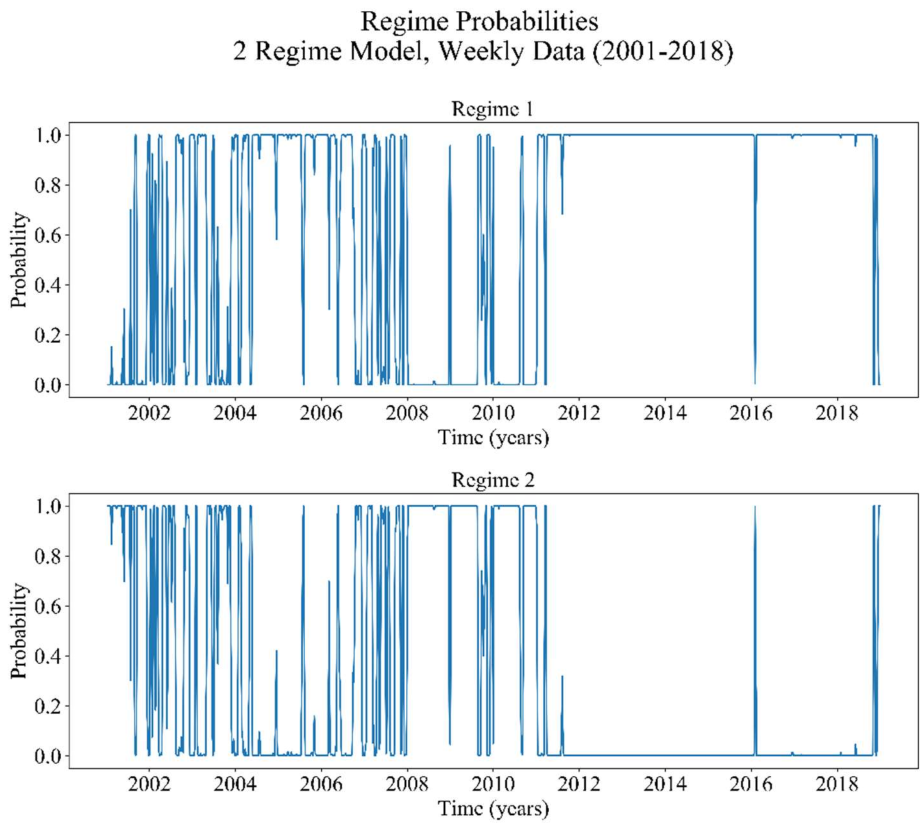


Figure 5.1.2: Regime posterior probabilities over time for a 2-regime model estimated using weekly data.



As increasing the number of regimes appeared to only decompose the period from 2008 to 2011 further, with the models remaining in Regime 1 for the majority of the period from late 2011 onwards, it was decided to only estimate model parameters using flow data from January 2012 onwards. This allowed for greater detail when examining the behaviour of portfolio flows from 2012 to 2016, avoiding scenarios where for example, a 4- or 5-regime model would only really have two regimes present after 2012.

Figure 5.1.3 and Figure 5.1.4 show 4-regime models estimated using daily and weekly flow data from January 2012 to December 2018. Regime switching behaviour, which was not observed in the previous models, is observed during the period 2012 to 2016. This is a positive result from the perspective of a regime-based asset allocation strategy, as being unable to observe any changes in the data for a 4 year period could lead to underperformance if, through inaction, the strategy failed to capture excess returns. It was also observed that regime switching was more frequent in the daily model compared to the weekly model. The daily model is estimated on a higher frequency and has more data available (1825 days vs 364 weeks) than the weekly model making it more sensitive to changes.

Figure 5.1.3: Regime posterior probabilities over time for a 4-regime model estimated using daily data from January 2012 onwards.

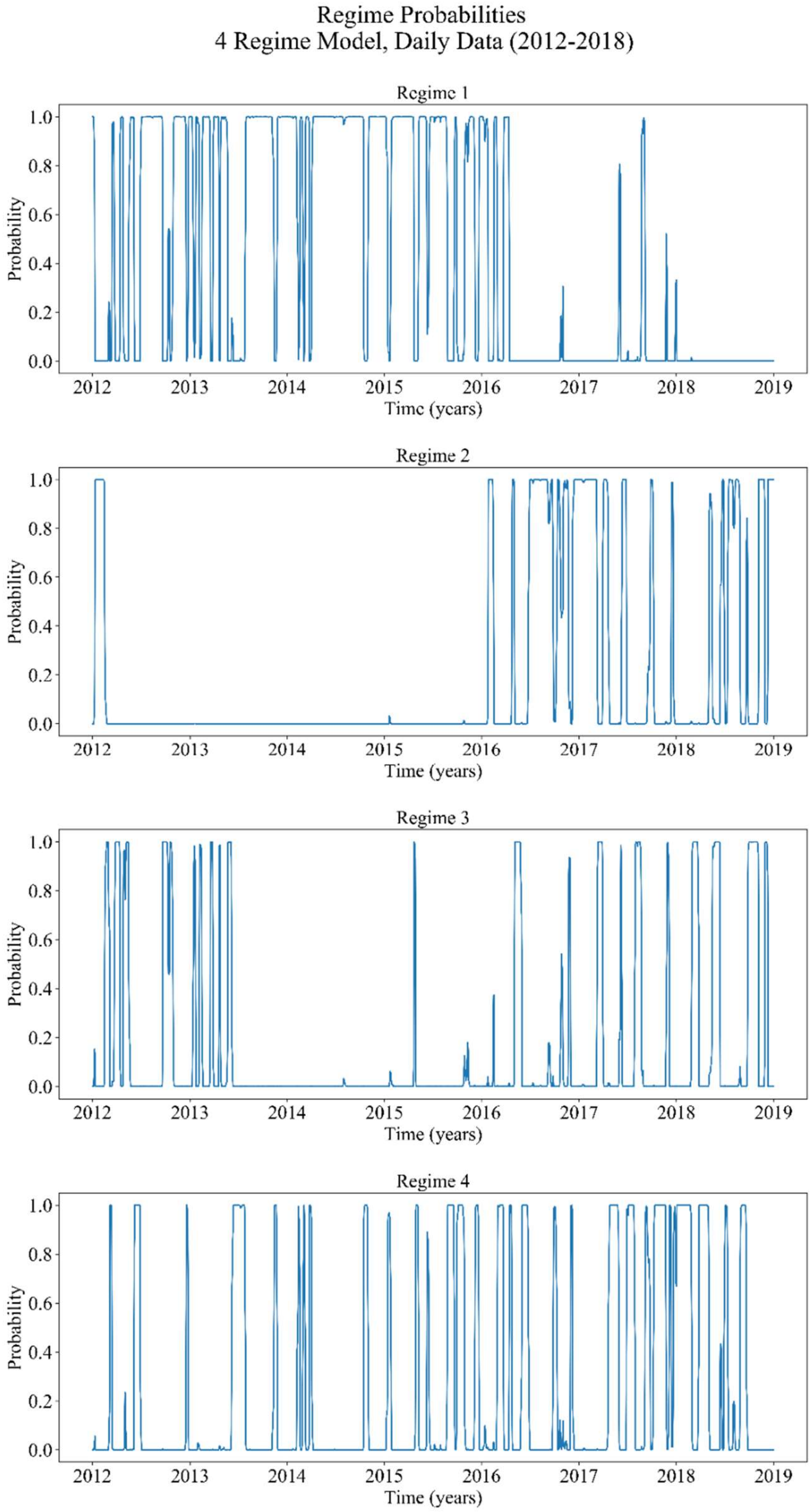
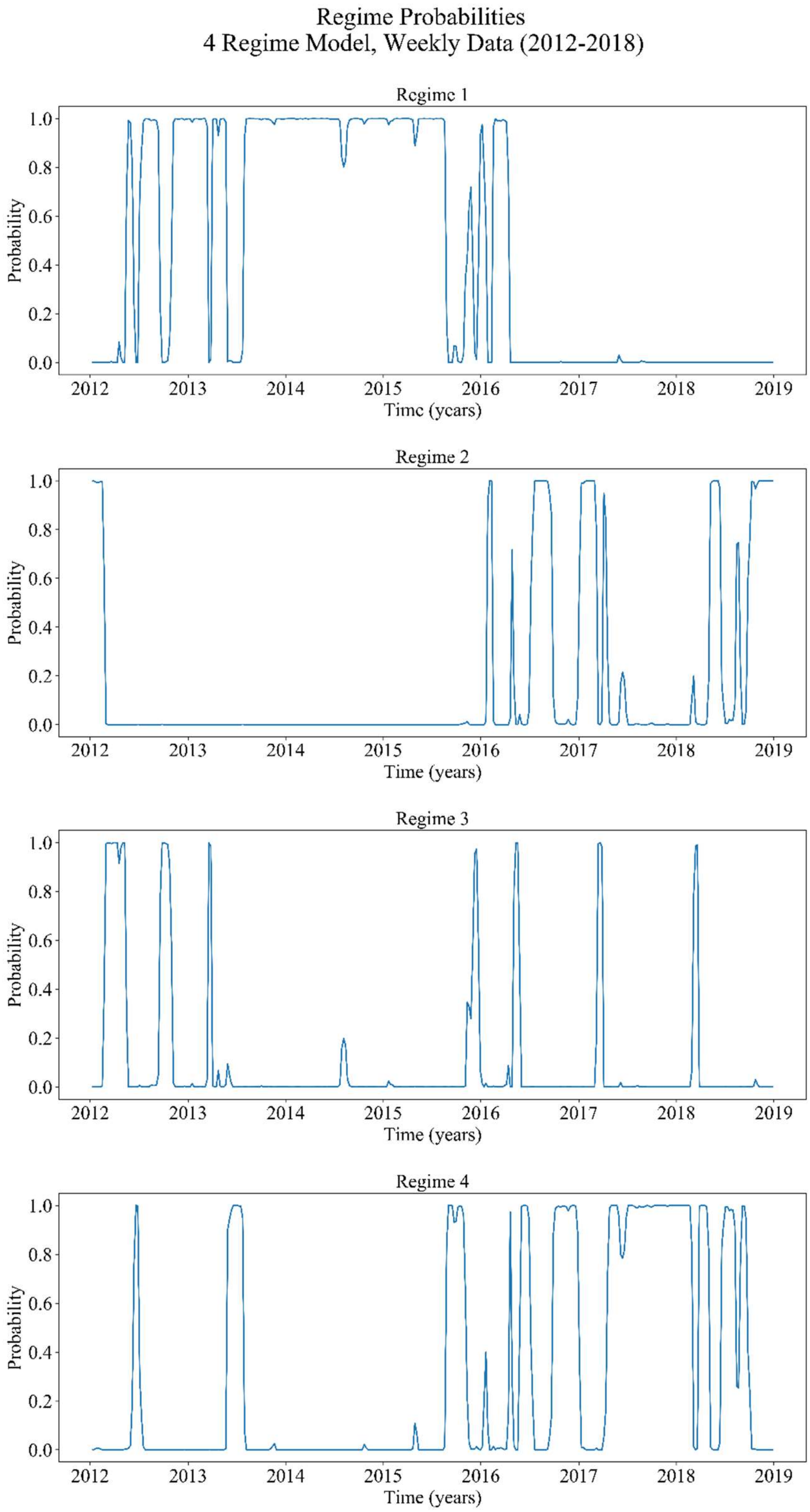


Figure 5.1.4: Regime posterior probabilities over time for a 4-regime model estimated using weekly data from January 2012 onwards.



5.1.3. Characteristics of Regimes

Regime-switching models were estimated using daily and weekly flow data from January 2012 to December 2018. The switching behaviour of these models is shown in Figure 5.1.3 and Figure 5.1.4. Descriptive statistics for this data, along with daily and weekly returns for the same period are shown in Table 4.3.2 (daily) and Table 4.3.3 (weekly). The parameters of the models estimated are shown in Table 5.1.5 and Table 5.1.7. The conditional distributions described in these tables show distinct values for each regime, separate from each other and from the original flow data. Descriptive statistics for daily and weekly returns during the periods when each regime was present in the model are shown in Table 5.1.6 and Table 5.1.8, respectively. Again, the values measured are distinct for each regime. This indicates that flow-based regimes exhibit characteristic returns, which could be used as part of a regime-switching asset allocation strategy. Results for 4-regime models are shown as they were found to be the best performing models in a series of robustness tests (Section 5.2.3).

A. *Daily Frequency*

Table 5.1.5: Parameters for a 4-regime Markov-switching model with Gaussian component distributions estimated using daily flow data from January 2012 – December 2018, where μ_{fi} and σ_{fi} parameterize the conditional distributions respectively.

Parameters for a 4-regime switching model estimated using daily flow data (Daily Data, 2012-2018)

	Regime 1		**Regime 2**		**Regime 3**		**Regime 4**	
	μ_{f1}	σ_{f1}	μ_{f2}	σ_{f2}	μ_{f3}	σ_{f3}	μ_{f4}	σ_{f4}
Europe ex UK	0.019	0.029	-0.026	0.093	0.020	0.062	0.036	0.082
EM Latin America	0.018	0.071	0.034	0.145	-0.156	0.147	0.105	0.141
EM EMEA	0.026	0.046	0.008	0.072	0.002	0.078	0.064	0.077
Pacific ex Japan	0.011	0.029	-0.038	0.067	-0.019	0.042	0.038	0.049
EM Asia	0.012	0.039	0.017	0.086	0.011	0.075	0.030	0.054
USA	-0.008	0.043	-0.133	0.081	-0.040	0.049	-0.061	0.085
UK	0.021	0.042	-0.098	0.164	-0.001	0.069	0.081	0.089
Japan	0.009	0.031	-0.048	0.060	-0.024	0.079	0.026	0.058

Table 5.1.6: Descriptive statistics for the daily returns associated with the 4-regime Markov-switching model with parameters described in Table 5.1.5, where μ_{ri} and σ_{ri} describe the mean returns and standard deviations for the daily returns in each regime respectively.

Descriptive statistics of daily returns (%) in each regime
(Daily Data, 2012-2018)

	Regime 1		Regime 2		Regime 3		Regime 4	
	μ_{r1}	σ_{r1}	μ_{r2}	σ_{r2}	μ_{r3}	σ_{r3}	μ_{r4}	σ_{r4}
Europe ex UK	0.035	0.986	0.028	0.919	-0.059	0.849	0.102	0.933
EM Latin America	-0.023	0.897	0.083	0.936	-0.050	0.873	0.122	1.014
EM EMEA	0.001	0.830	0.070	0.834	-0.045	0.808	0.099	0.904
Pacific ex Japan	0.003	0.723	0.064	0.636	-0.020	0.678	0.106	0.702
EM Asia	-0.013	0.760	0.060	0.724	-0.030	0.806	0.116	0.816
USA	0.027	0.793	0.051	0.773	-0.037	0.814	0.134	0.808
UK	0.009	0.839	0.063	0.843	-0.071	0.772	0.091	0.854
Japan	0.046	1.203	0.028	1.326	-0.105	1.099	0.176	1.221

B. *Weekly Frequency*

Table 5.1.7: Parameters for a 4-regime Markov-switching model with Gaussian component distributions estimated using weekly flow data from January 2012 – December 2018

Parameters for a 4-regime switching model estimated using daily flow data
(Weekly Data, 2012-2018)

	Regime 1		Regime 2		Regime 3		Regime 4	
	μ_{f1}	σ_{f1}	μ_{f2}	σ_{f2}	μ_{f3}	σ_{f3}	μ_{f4}	σ_{f4}
Europe ex UK	0.117	0.138	-0.171	0.434	0.182	0.236	0.110	0.385
EM Latin America	0.110	0.371	-0.272	0.893	-0.693	0.477	0.499	0.624
EM EMEA	0.165	0.234	-0.031	0.373	0.100	0.191	0.222	0.364
Pacific ex Japan	0.060	0.142	-0.210	0.328	-0.037	0.278	0.108	0.227
EM Asia	0.077	0.192	0.160	0.473	0.067	0.132	0.075	0.277
USA	-0.018	0.192	-0.525	0.343	-0.154	0.193	-0.459	0.421
UK	0.115	0.228	-0.597	0.724	0.023	0.275	0.324	0.471
Japan	0.070	0.148	-0.230	0.271	-0.091	0.303	0.011	0.322

Table 5.1.8: Descriptive statistics for the weekly returns associated with the 4-regime Markov-switching model with parameters described in Table 5.1.7, where μ_{ri} and σ_{ri} describe the mean returns and standard deviations for the weekly returns in each regime respectively

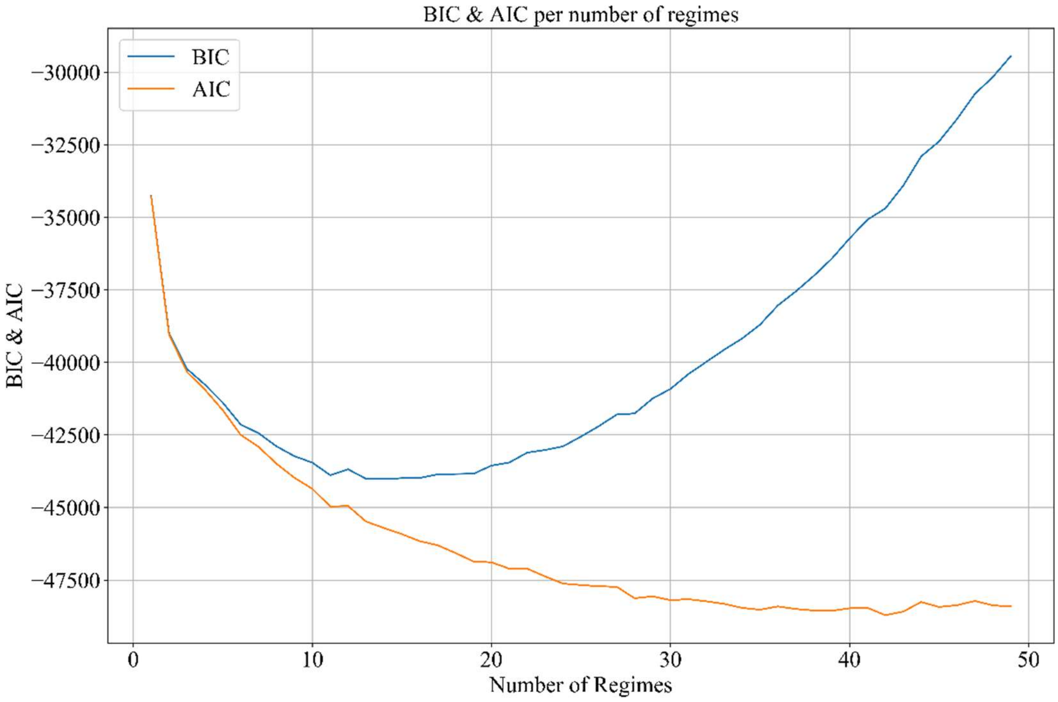
Descriptive statistics of weekly returns (%) in each regime (Weekly Data, 2012-2018)

	Regime 1		**Regime 2**		**Regime 3**		**Regime 4**	
	μ_{r1}	σ_{r1}	μ_{r2}	σ_{r2}	μ_{r3}	σ_{r3}	μ_{r3}	σ_{r3}
Europe ex UK	0.350	2.161	0.004	1.997	-0.282	1.876	0.113	1.911
EM Latin America	0.060	2.197	0.250	2.043	-0.304	1.537	0.251	1.963
EM EMEA	0.205	1.915	0.162	1.983	-0.372	1.718	0.178	1.915
Pacific ex Japan	0.209	1.651	0.173	1.315	0.036	1.677	0.111	1.634
EM Asia	0.097	1.863	0.398	1.788	-0.166	1.830	0.124	2.027
USA	0.303	1.631	0.095	1.972	-0.056	1.466	0.255	1.582
UK	0.186	1.862	0.138	1.715	-0.303	1.691	0.143	1.776
Japan	0.446	2.864	-0.162	2.766	-0.299	2.423	0.245	2.830

5.1.4. Model Selection

Figure 5.1.5 shows the BIC and AIC plotted as the number of regimes increases. As the number of possible regimes reaches 15, the BIC begins to increase, showing that this is the minimum score. The AIC continues to decrease as the penalty term used is smaller and takes less effect. The minimum score used is only found as the number of regimes increases past 30.

Figure 5.1.5: BIC & AIC calculated for an increasing number of regimes.



In theory, the number of regimes could be increased close to 15 based on the BIC without overfitting the model, however this is not practical in terms of the "explainability" of the model, also leading to stability issues when attempting to characterise the regimes based on returns.

The descriptive statistics for each regime are quite distinct, showing that an increased number of regimes has allowed the model to highlight more differences in the data. However, increasing the number of regimes also raises the issue of stability as the returns data used to calculate the descriptive statistics is divided into smaller and smaller portions for each regime. It was found in a series of robustness tests, that increasing the number of regimes past 4 decreases the stability of the model for characterising returns. This agrees with the finding of Gatumel and Ielpo (2014) who found that between two and five regimes were needed to model asset returns and Nystrup, Madsen et al. (2015) who found that four-regime models outperformed two- and three-regime models.

5.2. Testing a Regime-Based Asset Allocation Strategy

Using a portfolio of eight global indices a regime-based asset allocation strategy was tested, as described in Section 4.6. An out-of-sample walk-forward test was performed in an iterative fashion to obtain results which were as realistic as possible. Results are displayed below for a daily and weekly regime models. The main results displayed below are for 4-regime models, as this was deemed the best performing model in terms of stability in a series of robustness tests discussed in Section 5.2.3.

5.2.1. Daily Frequency

Two versions of the model were trained, with daily data up until time t used to determine the portfolio positions at time $t + \tau$. The first model has $\tau = 1$, with the predicted regime at time t determining portfolio positions 1 day ahead. The second model has $\tau = 3$, with the predicted state at time t determining the portfolio positions 3 days ahead. The performance of the regime-based strategy was compared to a static equal weight portfolio containing the same instruments as in the regime-based portfolio, as well as the MSCI World Total Return Index.

Table 5.2.1 shows key performance statistics - annualized return (AR), annualized volatility (Vol) and Sharpe Ratio (SR) - generated by the trading strategy in an out-of-sample test from January 2017 to December 2018.

The performance statistics are compared to a buy and hold strategy and to the MSCI World Total Return Index for the same period.

Table 5.2.1: Performances of strategies and indices from Jan 2017 to Dec 2018, daily data.

Strategy	AR (%)	Vol (%)	SR
Regime Model ($\tau = 1$)	16.23	7.55	1.84
Regime Model ($\tau = 3$)	14.97	7.46	1.69
Buy & Hold	5.89	8.07	0.45
MSCI World Index	4.77	9.75	0.25

The 1-day lag Regime Model realized the highest AR and SR. The 3-day lag Regime model had a slightly lower AR and SR but both strategies outperformed the benchmarks used for comparison across all performance indicators. The strong performance of the 3-day lag model demonstrates that returns captured using flow-based regimes are persistent, even if there is a delay in adjusting portfolio positions after the regime change.

The daily cumulative returns graphs show the ability of the regime models to identify periods where there is a change in regimes, and to adjust the long-short portfolio positions based on this to capture excess returns. Again, we see that the 1-day lag model (Figure 5.2.1) performs better than the 3-day lag model (Figure 5.2.2) as it is able to react more quickly to changes in market conditions, however the fact that the 3-day lag model is still able to capture excess returns shows that the changes in regime are relatively persistent.

Figure 5.2.1: 1-day lag regime model daily cumulative returns (January 2017 – December 2018)

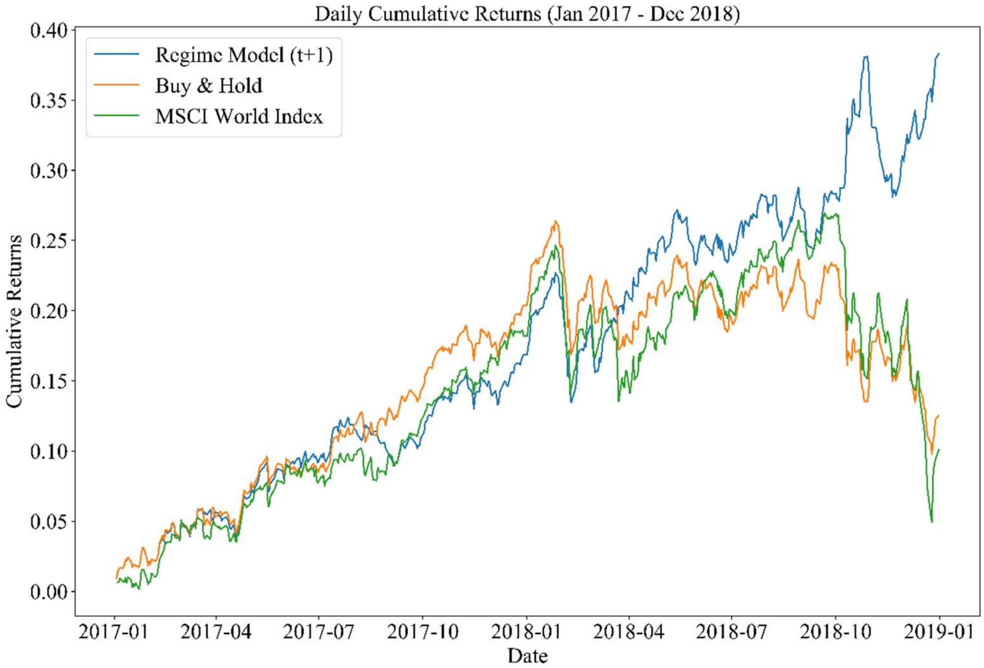
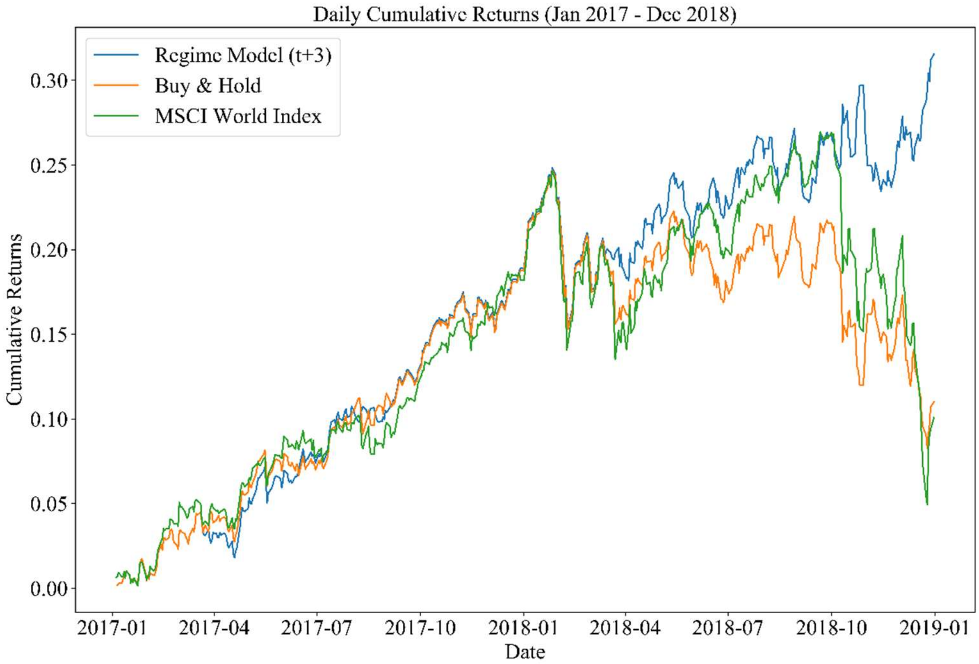


Figure 5.2.2: 3-day lag regime model daily cumulative returns (January 2017 – December 2018)



5.2.2. Weekly Frequency

The model was trained with weekly data up until time t used to determine the portfolio positions at time $t + \tau$. The model had $\tau = 1$, with the predicted regime at time t determining portfolio positions 1 week ahead. Models tested with $\tau > 1$ did not perform well. Intuitively, this makes sense as there would be at least a two-week delay on any action taken. The performance of the regime-based strategy was compared to a static equal weight portfolio containing the same instruments in the regime-based portfolio, as well as the MSCI World Total Return Index.

Table 5.2.2 shows key performance statistics - annualized return (AR), annualized volatility (Vol) and Sharpe Ratio (SR) - generated by the trading strategy in an out-of-sample test from January 2017 to December 2018.

The performance statistics are compared to a buy and hold strategy and to the MSCI World Total Return Index for the same period.

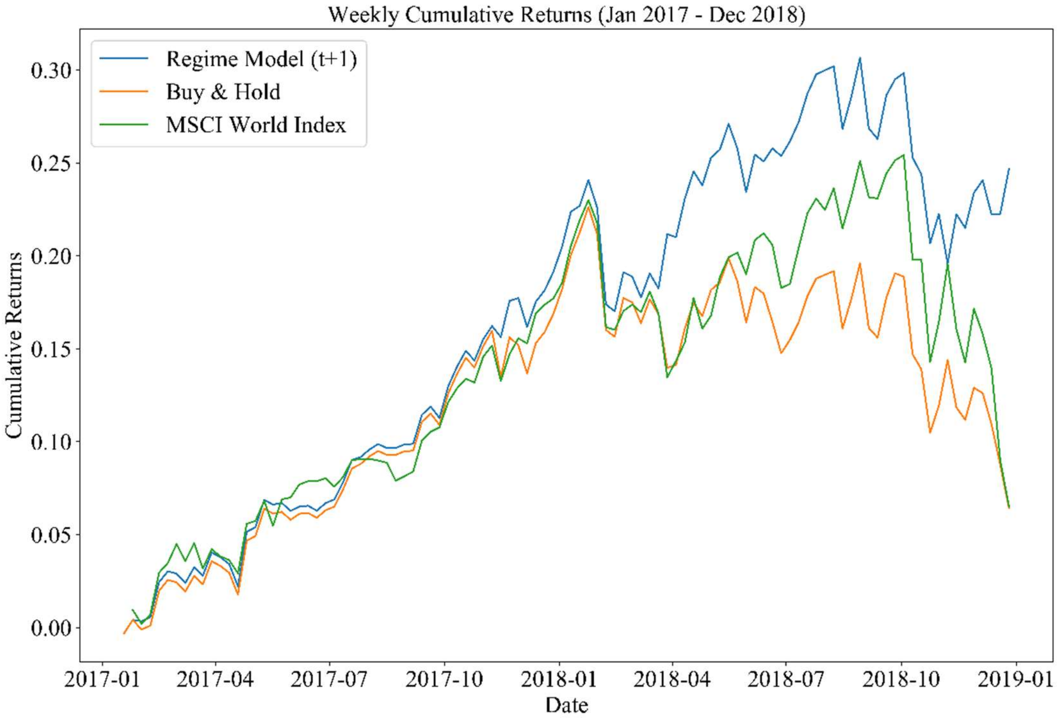
Table 5.2.2: Performances of strategies and indices from Jan 2017 to Dec 2018, weekly data.

Strategy	**AR (%)**	**Vol (%)**	**SR**
Regime Model ($\tau = 1$)	10.37	7.88	1.05
Buy & Hold	3.77	8.35	0.09
MSCI World Index	3.88	10.41	0.10

The weekly model had a lower AR and SR when compared to the daily models but still outperformed the benchmarks used for comparison across all performance indicators. This strong performance indicates again that flow-based regimes are persistent and can be used to capture excess returns on higher and lower time frames.

The cumulative returns graph in Figure 5.2.3 shows the ability of the weekly regime model to identify changes in regimes and that while the weekly model is slower to react to changes compared to the daily models, adjusting positions on a weekly basis outperforms a static buy and hold strategy.

Figure 5.2.3: 1-week lag regime model weekly cumulative returns (January 2017 – December 2018)



5.2.3. Robustness Tests

As the estimation of model parameters is an unsupervised process, there can be some variation in the regimes assigned to each period. To test the performance and stability of the daily and weekly models over a number of regimes, the model testing and training was repeated a number of times for each regime to determine the stability of the results. The results of these stability tests are given as a 95% confidence interval for the annualized return of the model.

The average annualized returns for the daily and weekly models, and the confidence intervals for those figures are shown in Table 5.2.3, Table 5.2.4 and Table 5.2.5. LIBOR was used as the risk-free rate when calculating the Sharpe Ratio. The results are compared to a Buy & Hold strategy over the same period, as well as the MSCI Total Return World Index.

For both daily and weekly results, each model was re-tested 100 times. The 4-regime model was the best performing model, both in terms of returns and standard deviation. Both daily and weekly models outperformed the Buy & Hold and MSCI World Index in terms of annualized returns and volatility, with the 4-regime model being the highest performing in terms of returns and stability on both daily and weekly frequencies. As the number of regimes increases, the confidence interval also increases, meaning that there is added uncertainty in each model's predictions.

Table 5.2.3: Daily model $(\tau = 1)$ *performance (January 2017 – December 2018, 100 tests)*

	4 Regime Model	**5 Regime Model**	**6 Regime Model**	**7 Regime Model**	**Buy & Hold**	**MSCI World Index**
Annualized Returns (%)	**16.23**	**15.53**	**16.88**	**16.22**	5.89	4.77
95% Confidence Interval (±%)	1.88	1.89	2.41	3.04	-	-
Volatility (%)	7.55	7.12	7.31	7.15	8.07	9.75
Sharpe Ratio	1.84	1.84	2.00	1.89	-	-

Table 5.2.4: Daily model ($\tau = 3$) performance (January 2017 – December 2018, 100 tests)

	4 Regime Model	5 Regime Model	6 Regime Model	7 Regime Model	Buy & Hold	MSCI World Index
Annualized Returns (%)	14.97	10.58	12.39	11.18	5.89	4.77
95% Confidence Interval (±%)	2.06	2.30	2.12	3.53	-	-
Volatility (%)	7.46	7.46	7.62	7.34	8.07	9.75
Sharpe Ratio	1.69	1.84	1.37	1.24	0.45	0.25

Table 5.2.5: Weekly model ($\tau = 1$) performance (January 2017 – December 2018, 100 tests)

	4 Regime Model	5 Regime Model	6 Regime Model	7 Regime Model	Buy & Hold	MSCI World Index
Annualized Returns (%)	10.37	9.64	11.40	9.93	3.77	3.88
95% Confidence Interval (±%)	1.79	2.60	2.84	3.40	-	-
Volatility (%)	7.88	7.93	7.92	7.65	8.35	10.41
Sharpe Ratio	1.05	0.97	1.20	1.03	0.9	0.10

Chapter 6. Conclusion

This thesis was focused around two main research objectives. The first objective was to explore the growing applications of machine learning in investing, specifically within the areas of quantitative finance.

A review of the academic literature demonstrated that machine learning enables deeper financial analysis and research partly by allowing for the modelling of complex non-linear relationships in financial data. Machine learning also provides a means to analyse new forms of alternative data such as news, earnings calls and social media activity. There is also potential to innovate backtesting and strategy verification methodologies, through the development of new machine learning-based testing methods and performance metrics. The development and continuous improvement of programming languages and frameworks has allowed a wide range of algorithms and approaches to be used when applying machine learning across various areas within finance.

The above findings are valuable as they provide an understanding of how the use of machine learning has developed within quantitative finance. This could prove useful to both researchers and investors looking to apply machine learning within their own work, as well as providing a solid foundation for the second research objective in this thesis.

The second research objective was to demonstrate how machine learning can be used in the investment process to extract information from international portfolio flow data. The approach chosen was to construct regime-switching models, a machine learning method which has been successfully used to model and capture the behaviour of returns. The parameters of hidden Markov models were estimated using portfolio flow data from eight global regions. Examining the estimated parameters showed that the models were able to capture characteristics of the flow data which had previously been found in the literature. The values of the transition probability matrix diagonals confirmed that the models had captured the persistence of portfolio flows compared to returns. It was possible to characterise each regime based on their returns, by examining periods when certain regimes were present over time. Each regime was shown to have unique characteristic returns, in line with previous literature which found a relationship between flows and equity returns.

By examining the switching behaviour of the models, certain persistent regime shifts and structural breaks were found. Flow data describes the behaviour of informed investors, and these structural breaks were consistent across models with varying numbers of regimes, and on both daily and weekly frequencies. It could be possible that regime-switching models trained using flow data capture fundamental shifts in investor behaviour, for example due to the Great Financial Crisis, or quantitative easing policies. However, further research would be required to confirm this.

Based on these findings, a proof-of-concept regime-based asset allocation strategy was constructed, with hidden Markov models estimated using flow data being used to define regimes, which were then characterised based on their mean returns. In an out-of-sample walk-forward test, long or short positions were adjusted in eight MSCI country and regional indices based on the mean return characteristics of the current regime determined by the hidden Markov model. It was found that the regime-based asset allocation strategy significantly outperformed a buy and hold strategy as well as the MSCI World Index over the test period across all performance metrics used. This finding held true on models trained on both daily and weekly frequencies. It was also found that 4-regime models gave the most stable performance during the test period in a series of robustness tests over daily and weekly frequency.

By using hidden Markov models, an unsupervised machine learning technique, it was possible to extract information from international portfolio flow data, confirming certain properties of the data found in the literature. Using this information to construct a profitable proof-of-concept machine learning based trading strategy demonstrates how this approach can aid and improve the investment process.

There are several areas where this research could be continued. The appearance of structural breaks and persistent regime shifts which occur seemingly in conjunction with periods of notable economic and financial change warrants further research. As flows capture the behaviour of informed investors, further research could examine whether regime-switching models can capture investor behaviour as during crisis situations or during policy changes. While this research focused on equity flows, including data from other asset classes could also prove useful in capturing investor behaviour. Most dynamic asset-allocation strategies focus on the adjustment of allocations between equities, bonds, and cash. A regime-switching model which included flows in bonds, currencies and other asset classes would give greater insight into investor behaviour and could result in a more accurate model. Similarly, the parameters of the hidden Markov models used in this research assumed Gaussian distributions. Other distributions such as student-t could potentially provide a better fit for the data.

Both research objectives were achieved in this thesis, the literature review provides an overview of the applications of machine learning in finance, while the regime-switching model demonstrates a practical application of machine learning for extracting information from financial data. Taken as a whole, the research explores how machine learning has developed as a tool for financial analysis and demonstrates its usefulness by capturing key information from a financial dataset and using it to improve the investment process.

Bibliography

(2019). State Street Corporation 10-K for Fiscal Year ended Dec 31, 2018. EDGAR, Securities and Exchange Commission.

Abe, M. and H. Nakayama (2018). Deep Learning for Forecasting Stock Returns in the Cross-Section. Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer.

Achab, M., et al. (2018). Profitable Bandits. Proceedings of The 10th Asian Conference on Machine Learning. Z. Jun and T. Ichiro. Proceedings of Machine Learning Research, PMLR. **95:** 694-709.

Al-Aradi, A. and S. Jaimungal (2019). "Active and Passive Portfolio Management with Latent Factors." arXiv preprint arXiv:1903.06928.

Alberg, J. and Z. C. Lipton (2017). "Improving Factor-Based Quantitative Investing by Forecasting Company Fundamentals." arXiv preprint arXiv:1711.04837.

Ammann, M. and M. Verhofen (2006). "The effect of market regimes on style allocation." Financial Markets and Portfolio Management **20**(3): 309-337.

Ang, A. and G. Bekaert (2002). "International Asset Allocation With Regime Shifts." The Review of Financial Studies **15**(4): 1137-1187.

Ang, A. and G. Bekaert (2004). "How Regimes Affect Asset Allocation." Financial Analysts Journal **60**(2): 86-99.

Ang, A. and A. Timmermann (2012). "Regime Changes and Financial Markets." Annual Review of Financial Economics **4**(1): 313-337.

Arnott, R., et al. (2019). "A Backtesting Protocol in the Era of Machine Learning." The Journal of Financial Data Science **1**(1): 64-74.

Barga, R., et al. (2015). Introducing Microsoft Azure Machine Learning. Predictive Analytics with Microsoft Azure Machine Learning, Springer: 21-43.

Bauer, R., et al. (2004). "Asset allocation in stable and unstable times." The Journal of Investing **13**(3): 72-80.

Baum, L. E., et al. (1970). "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains." The annals of mathematical statistics **41**(1): 164-171.

Becker, Y. L. and M. R. Reinganum (2018). The Current State of Quantitative Equity Investing, CFA Institute Research Foundation.

Belloni, A., et al. (2014). "Inference on treatment effects after selection among high-dimensional controls." The Review of Economic Studies **81**(2): 608-650.

Bianchi, D., et al. (2018). "Bond risk premia with machine learning." Available at SSRN 3232721.

Brennan, M. J. and H. H. Cao (1997). "International portfolio investment flows." The journal of Finance **52**(5): 1851-1880.

Buehler, H., et al. (2019). "Deep hedging." Quantitative Finance: 1-21.

Bulla, J. and I. Bulla (2006). "Stylized facts of financial time series and hidden semi-Markov models." Computational Statistics & Data Analysis **51**(4): 2192-2209.

Bulla, J., et al. (2011). "Markov-switching asset allocation: Do profitable strategies exist?" Journal of Asset Management **12**(5): 310-321.

Carhart, M. M. (1997). "On persistence in mutual fund performance." The journal of Finance **52**(1): 57-82.

Chow, J. C. (2018). "Analysis of Financial Credit Risk Using Machine Learning." arXiv preprint arXiv:1802.05326.

Choy, G., et al. (2018). "Current applications and future impact of machine learning in radiology." Radiology **288**(2): 318-328.

Cochrane, J. H. (2011). "Presidential Address: Discount Rates." The journal of Finance **66**(4): 1047-1108.

Cont, R. (2001). "Empirical properties of asset returns: stylized facts and statistical issues."

De Prado, M. L. (2019). "Tactical Investment Algorithms." Available at SSRN: https://ssrn.com/abstract=3459866.

De Spiegeleer, J., et al. (2018). "Machine learning for quantitative finance: fast derivative pricing, hedging and fitting." Quantitative Finance: 1-9.

Deng, L. and D. Yu (2014). "Deep learning: methods and applications." Foundations and Trends® in Signal Processing **7**(3–4): 197-387.

Dixon, M., et al. (2017). "Classification-based financial markets prediction using deep neural networks." Algorithmic Finance(Preprint): 1-11.

Domingos, P. (2012). "A few useful things to know about machine learning." Communications of the ACM **55**(10): 78-87.

Fama, E. F. and K. R. French (1993). "Common risk factors in the returns on stocks and bonds." Journal of financial economics **33**(1): 3-56.

Fama, E. F. and J. D. MacBeth (1973). "Risk, return, and equilibrium: Empirical tests." Journal of political economy **81**(3): 607-636.

Fecamp, S., et al. (2019). "Risk management with machine-learning-based algorithms." arXiv preprint arXiv:1902.05287.

Feng, G., et al. (2017). "Taming the factor zoo."

Ferrucci, D., et al. (2010). "Building Watson: An overview of the DeepQA project." AI magazine **31**(3): 59-79.

Fons, E., et al. (2019). "A novel dynamic asset allocation system using Feature Saliency Hidden Markov models for smart beta investing." arXiv preprint arXiv:1902.10849.

French, C. W. (2003). "The Treynor capital asset pricing model." Journal of Investment Management **1**(2): 60-72.

Froot, K. and M. Teo (2008). "Style investing and institutional investors." Journal of Financial and Quantitative Analysis **43**(4): 883-906.

Froot, K. A., et al. (2014). "Multi-Asset Sentiment and InstitutionalInvestor Behavior: A Cross-Asset Perspective." The Journal of Portfolio Management **40**(4): 144-156.

Froot, K. A., et al. (2001). "The portfolio flows of international investors." Journal of financial economics **59**(2): 151-193.

Froot, K. A. and T. Ramadorai (2008). "Institutional Portfolio Flows and International Investments." The Review of Financial Studies **21**(2): 937-971.

Gatumel, M. and F. Ielpo (2014). "The number of regimes across asset returns: identification and economic value." International Journal of Theoretical and Applied Finance **17**(06): 1450040.

Gonzalvez, J., et al. (2019). "Financial Applications of Gaussian Processes and Bayesian Optimization." arXiv preprint arXiv:1903.04841.

Goodfellow, I., et al. (2014). Generative adversarial nets. Advances in neural information processing systems.

Goudenège, L., et al. (2019). "Gaussian Process Regression for Pricing Variable Annuities with Stochastic Volatility and Interest Rate." arXiv preprint arXiv:1903.00369.

Graham, B. and D. Dodd (2008). Security Analysis: Foreword by Warren Buffett, McGraw-Hill Professional.

Griffin, J. M., et al. (2004). "Are daily cross-border equity flows pushed or pulled?" Review of Economics and Statistics **86**(3): 641-657.

Guidolin, M. (2011). Markov switching models in empirical finance. Missing data methods: Time-series methods and applications, Emerald Group Publishing Limited**:** 1-86.

Guidolin, M. and A. Timmermann (2007). "Asset allocation under multivariate regime switching." Journal of Economic Dynamics and Control **31**(11): 3503-3544.

Guidolin, M. and A. Timmermann (2008). "International asset allocation under regime switching, skew, and kurtosis preferences." The Review of Financial Studies **21**(2): 889-935.

Hamilton, J. D. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle." Econometrica: Journal of the econometric society: 357-384.

Hamilton, J. D. and R. Susmel (1994). "Autoregressive conditional heteroskedasticity and changes in regime." Journal of econometrics **64**(1-2): 307-333.

Harrington, P. (2012). "Machine learning in action." Shelter Island, NY: Manning Publications Co.

Harvey, C. and Y. Liu (2017). "Lucky factors."

Harvey, C. R., et al. (2016). "… and the Cross-Section of Expected Returns." The Review of Financial Studies **29**(1): 5-68.

Heaton, J., et al. (2017). "Deep learning for finance: deep portfolios." Applied Stochastic Models in Business and Industry **33**(1): 3-12.

Hess, M. K. (2006). "Timing and diversification: A state-dependent asset allocation approach." European Journal of Finance **12**(03): 189-204.

Hisano, R., et al. (2018). "Predicting Adverse Media Risk using a Heterogeneous Information Network." arXiv preprint arXiv:1811.12166.

Jiang, Z., et al. (2017). "A deep reinforcement learning framework for the financial portfolio management problem." arXiv preprint arXiv:1706.10059.

Kahn, R. N. (2018). The Future of Investment Management, CFA Institute Research Foundation.

Kakushadze, Z. and W. Yu (2016). "Statistical industry classification." Journal of Risk & Control **3**(1): 17-65.

Kakushadze, Z. and W. Yu (2017). "Statistical risk models." The Journal of Investment Strategies **6**(2): 1-40.

Kakushadze, Z. and W. Yu (2019). "Machine Learning Risk Models." Journal of Risk & Control **6**(1): 37-64.

Khaidem, L., et al. (2016). "Predicting the direction of stock market prices using random forest." arXiv preprint arXiv:1605.00003.

Kinn, D. (2018). "Reducing Estimation Risk in Mean-Variance Portfolios with Machine Learning." arXiv preprint arXiv:1804.01764.

Kritzman, M., et al. (2012). "Regime Shifts: Implications for Dynamic Strategies (corrected)." Financial Analysts Journal **68**(3): 22-39.

Le, L., et al. (2015). On predictability of rare events leveraging social media: a machine learning perspective. Proceedings of the 2015 ACM on Conference on Online Social Networks, ACM.

Liang, Z., et al. (2018). "Deep Reinforcement Learning in Portfolio Management." arXiv preprint arXiv:1808.09940.

Lindsay, R. K., et al. (1993). "DENDRAL: a case study of the first expert system for scientific hypothesis formation." Artificial intelligence **61**(2): 209-261.

Lintner, J. (1975). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Stochastic Optimization Models in Finance, Elsevier**:** 131-155.

Liu, W., et al. (2017). "A survey of deep neural network architectures and their applications." Neurocomputing **234**: 11-26.

Lopez de Prado, M. (2016). "Building diversified portfolios that outperform out-of-sample." Journal of Portfolio Management.

Lopez de Prado, M. (2016). "Mathematics and economics: a reality check."

Lopez de Prado, M. and M. J. Lewis (2018). "Detection of False Investment Strategies Using Unsupervised Learning Methods." Available at SSRN: https://ssrn.com/abstract=3167017 or http://dx.doi.org/10.2139/ssrn.3167017.

Maheu, J. M., et al. (2012). "Components of bull and bear markets: bull corrections and bear rallies." Journal of Business & Economic Statistics **30**(3): 391-403.

Malkiel, B. G. and E. F. Fama (1970). "Efficient capital markets: A review of theory and empirical work." The journal of Finance **25**(2): 383-417.

Markowitz, H. (1952). "Portfolio selection." The journal of Finance **7**(1): 77-91.

McCulloch, W. S. and W. Pitts (1943). "A logical calculus of the ideas immanent in nervous activity." The bulletin of mathematical biophysics **5**(4): 115-133.

Modha, D. and C. Witchalls (2014). A computer that thinks, New Scientist Ltd. **224:** 28-29.

Mossin, J. (1966). "Equilibrium in a capital asset market." Econometrica: Journal of the econometric society: 768-783.

Nakagawa, K., et al. (2019). Deep Recurrent Factor Model: Interpretable Non-Linear and Time-Varying Multi-Factor Model. AAAI-19 Workshop on Network Interpretability for Deep Learning, Honolulu, Hawaii, USA, arXiv preprint arXiv:1901.11493.

Nakagawa, K., et al. (2018). "Deep Factor Model." arXiv preprint arXiv:1810.01278.

Nakagawa, K., et al. (2018). Deep factor model. ECML PKDD 2018 Workshops, Springer.

Nousi, P., et al. (2018). "Machine learning for forecasting mid price movement using limit order book data." arXiv preprint arXiv:1809.07861.

Nystrup, P., et al. (2017). "Dynamic Allocation or Diversification: A Regime-Based Approach to Multiple Assets." The Journal of Portfolio Management **44**(2): 62-73.

Nystrup, P., et al. (2015). "Regime-Based Versus Static Asset Allocation: Letting the Data Speak." The Journal of Portfolio Management **42**(1): 103-109.

Nystrup, P., et al. (2015). "Stylised facts of financial time series and hidden Markov models in continuous time." Quantitative Finance **15**(9): 1531-1541.

Passalis, N., et al. (2019). "Temporal Logistic Neural Bag-of-Features for Financial Time series Forecasting leveraging Limit Order Book Data." arXiv preprint arXiv:1901.08280.

Procacci, P. F. and T. Aste (2018). "Forecasting market states." Available at SSRN: https://ssrn.com/abstract=3215945 or http://dx.doi.org/10.2139/ssrn.3215945.

Raffinot, T. (2017). "Hierarchical clustering based asset allocation." Available at SSRN 2840729.

Rana, R. and F. S. Oliveira (2015). "Dynamic pricing policies for interdependent perishable products or services using reinforcement learning." Expert systems with applications **42**(1): 426-436.

Recht, B. and A. Rahimi (2017). Reflections on random kitchen sinks, 2017.

Reynolds, H. (2017). "AI? Or cognitive computing?" KM World **26**(9): 4-5.

Richard, J.-C. and T. Roncalli (2019). "Constrained Risk Budgeting Portfolios: Theory, Algorithms, Applications & Puzzles." arXiv preprint arXiv:1902.05710.

Ritter, G. (2017). "Machine learning for trading."

Rosenberg, B. (1974). "Extra-market components of covariance in security returns." Journal of Financial and Quantitative Analysis **9**(2): 263-274.

Ross, S. A. (2013). The arbitrage theory of capital asset pricing. HANDBOOK OF THE FUNDAMENTALS OF FINANCIAL DECISION MAKING: Part I, World Scientific**:** 11-30.

Russell, S. and P. Norvig (2009). Artificial Intelligence A Modern Approach Third Edition, Prentice Hall.

Rydén, T., et al. (1998). "Stylized facts of daily return series and the hidden Markov model." Journal of applied econometrics **13**(3): 217-244.

Samo, Y.-L. K. and A. Vervuurt (2016). Stochastic Portfolio Theory: a machine learning perspective. Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence. Jersey City, New Jersey, USA, AUAI Press: 657-665.

Sardelich, M. and S. Manandhar (2018). "Multimodal deep learning for short-term stock volatility prediction." arXiv preprint arXiv:1812.10479.

Sculley, D., et al. (2018). "Winner's Curse? On Pace, Progress, and Empirical Rigor."

Shah, D., et al. (2018). Predicting the Effects of News Sentiments on the Stock Market. 2018 IEEE International Conference on Big Data (Big Data), IEEE.

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

Song, Y.-G., et al. (2018). "Neural networks for stock price prediction." arXiv preprint arXiv:1805.11317.

Taghizadeh, E. (2017). "Utilizing artificial neural networks to predict demand for weather-sensitive products at retail stores." arXiv preprint arXiv:1711.08325.

Tesar, L. L. and I. M. Werner (1994). International equity transactions and US portfolio choice. The internationalization of equity markets, University of Chicago Press: 185-227.

Tesar, L. L. and I. M. Werner (1995). "Home bias and high turnover." Journal of international Money and Finance **14**(4): 467-492.

Tsantekidis, A., et al. (2017). Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks. 2017 IEEE 19th Conference on Business Informatics (CBI).

Viterbi, A. (1967). "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm." IEEE transactions on Information Theory **13**(2): 260-269.

Wagner, W. P. (2017). "Trends in expert system development: A longitudinal content analysis of over thirty years of expert system case studies." Expert systems with applications **76**: 85-96.

Wang, Y. and X. S. Ni (2019). "A XGBoost risk model via feature selection and Bayesian hyper-parameter optimization." International Journal of Database Management Systems (IJDMS) preprint arXiv:1901.08433.

Wiecki, T., et al. (2016). "All That Glitters Is Not Gold: Comparing Backtest and Out-of-Sample Performance on a Large Cohort of Trading Algorithms." The Journal of Investing **25**(3): 69-80.

Witten, I. H., et al. (2016). Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann.

Zhang, M., et al. (2018). "Latent Dirichlet Allocation with Residual Convolutional Neural Network Applied in Evaluating Credibility of Chinese Listed Companies." arXiv preprint arXiv:1811.11017.

Appendix 1: Model Switching Behaviour

This appendix contains figures showing the models posterior probabilities over time - the probability that a regime was present at time t - for models estimated using daily and weekly flow data over the period January 2001 to December 2018.

Section 5.1.2 contains figures describing the switching behaviour of 2-regime models from January 2001 to December 2018, where a structural break appears in the flow data around the end of 2011. After this point certain regimes no longer appear to be present and a certain regime appears to dominate the switching behaviour. This appendix contains additional figures for 3-, 4-, and 5-regime models, demonstrating similar switching behaviour over the same period

When a 3-regime model is estimated over the same period the model again remains in Regime 1 for a long period from late 2011 to 2016, with Regime 2 being the dominant period from 2008 to 2011 (Figure A.1.1, Figure A.1.2)

The 4-regime and 5-regime models also appear to exhibit similar behaviour, remaining in Regime 1 for a long period from late 2011 to 2016. However, the period from 2008 to 2011 is further decomposed into Regimes 2 and 3 in the 4-regime model (Figure A.1.3 & Figure A.1.4), and Regimes 2, 3 and 4 in the 5-regime model (Figure A.1.5 & Figure A.1.6).

As discussed in Section 5.1.2, a possible economic explanation that could be assigned to the persistent presence of certain regimes during period 2008 to 2011, is the Global Financial Crisis. The period after this from 2011 to 2016 where there is an apparent structural break, could possibly be related to factors such as monetary policy and quantitative easing after the Financial Crisis. These, or other possible explanations for regime switching behaviour were not verified or explored further but could provide an interesting topic for further research.

Figure A.1.1: Regime posterior probabilities over time for a 3-regime model estimated using daily data.

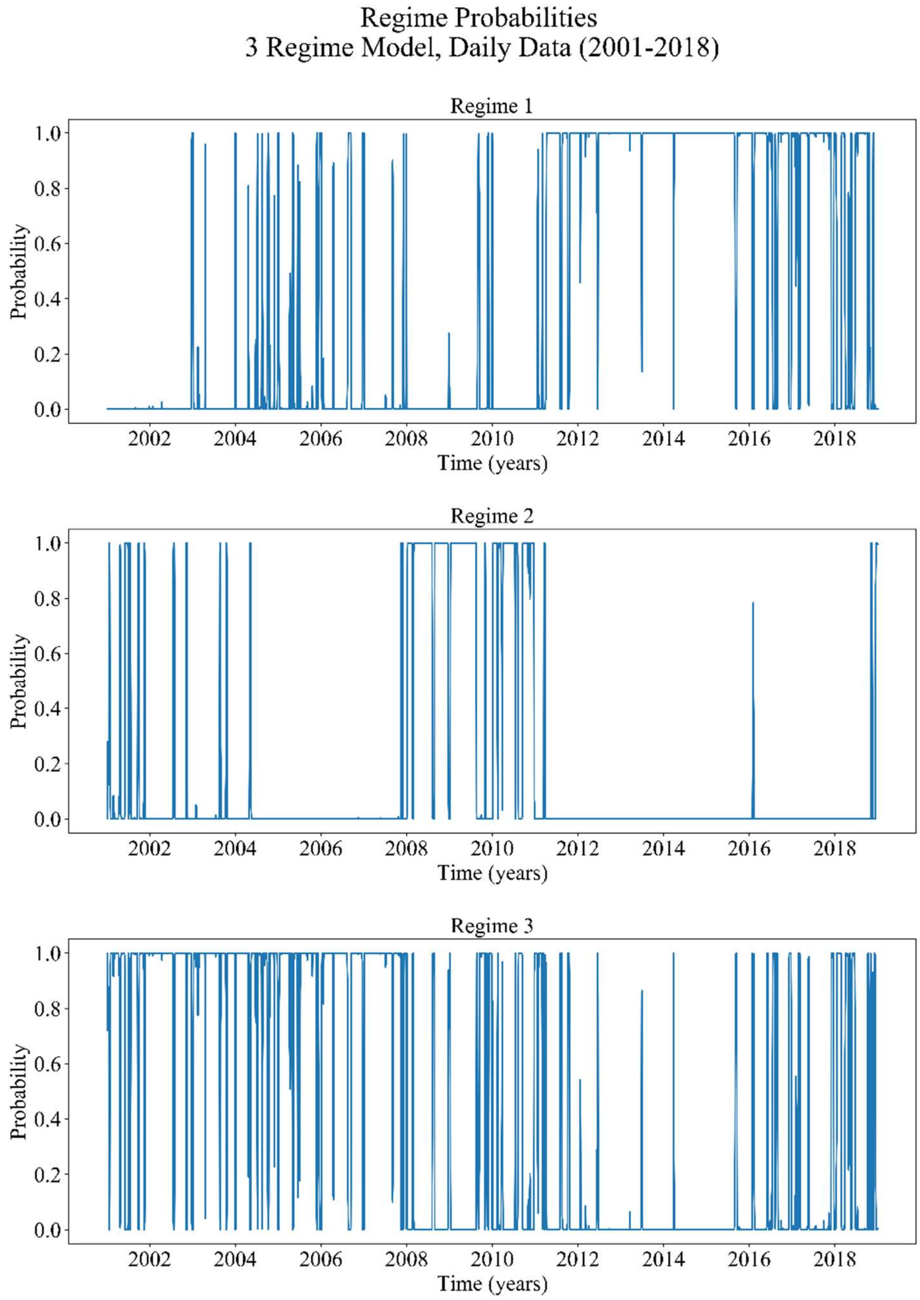


Figure A.1.2: Regime posterior probabilities over time for a 3-regime model estimated using weekly data.

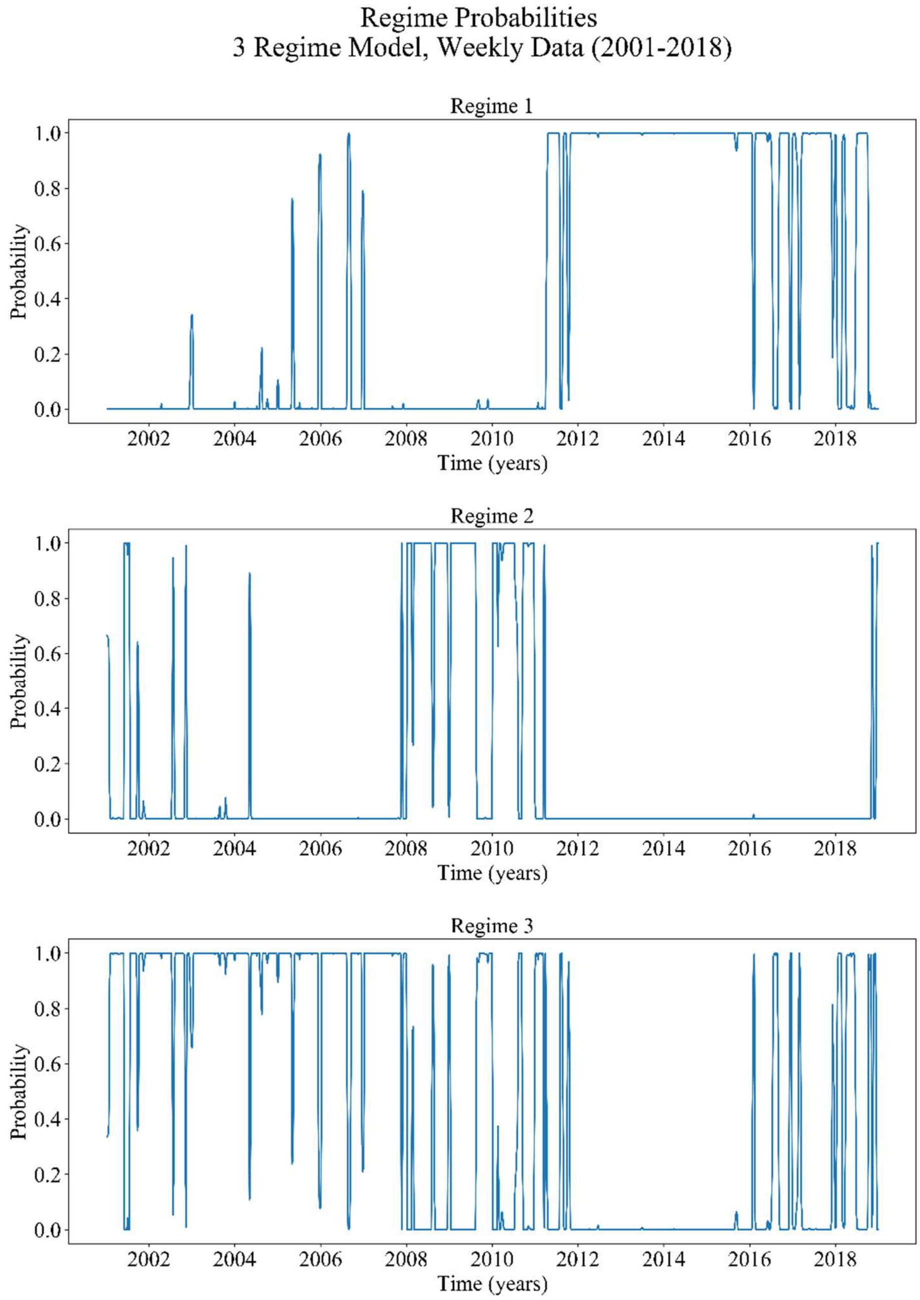


Figure A.1.3: Regime posterior probabilities over time for a 4-regime model estimated using daily data.

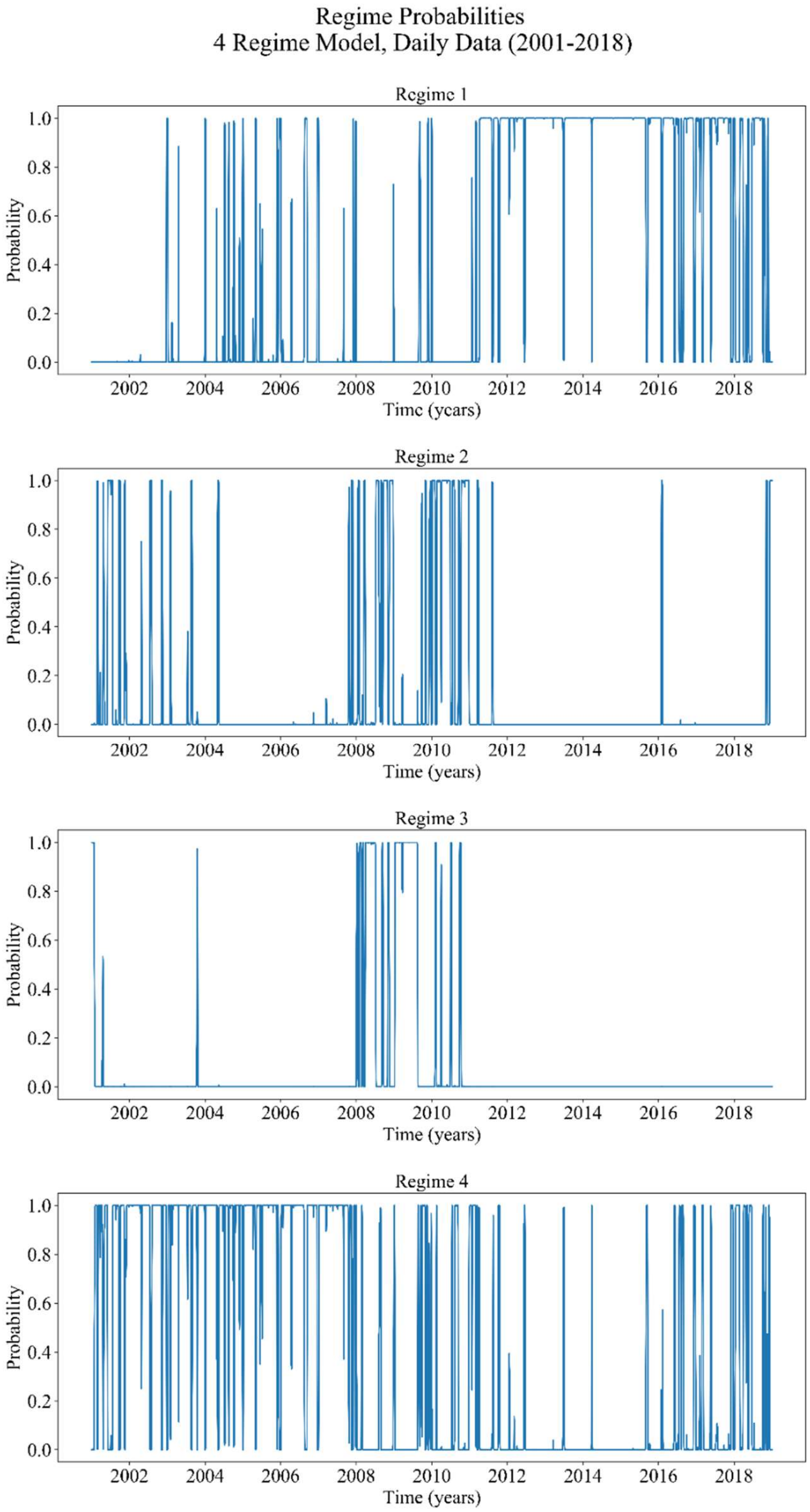


Figure A.1.4: Regime posterior probabilities over time for a 4-regime model estimated using weekly data.

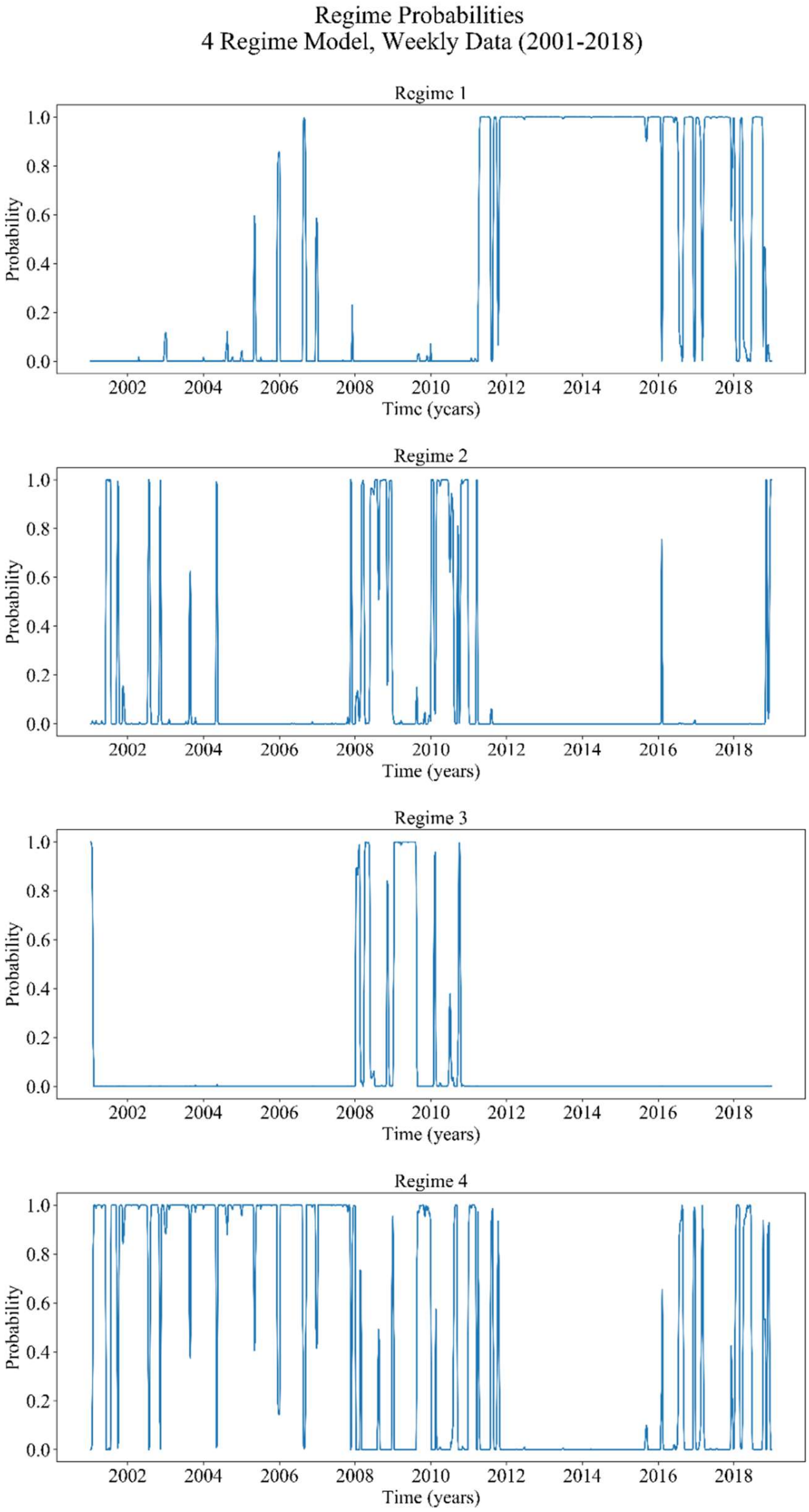


Figure A.1.5: Regime posterior probabilities over time for a 5-regime model estimated using daily data.

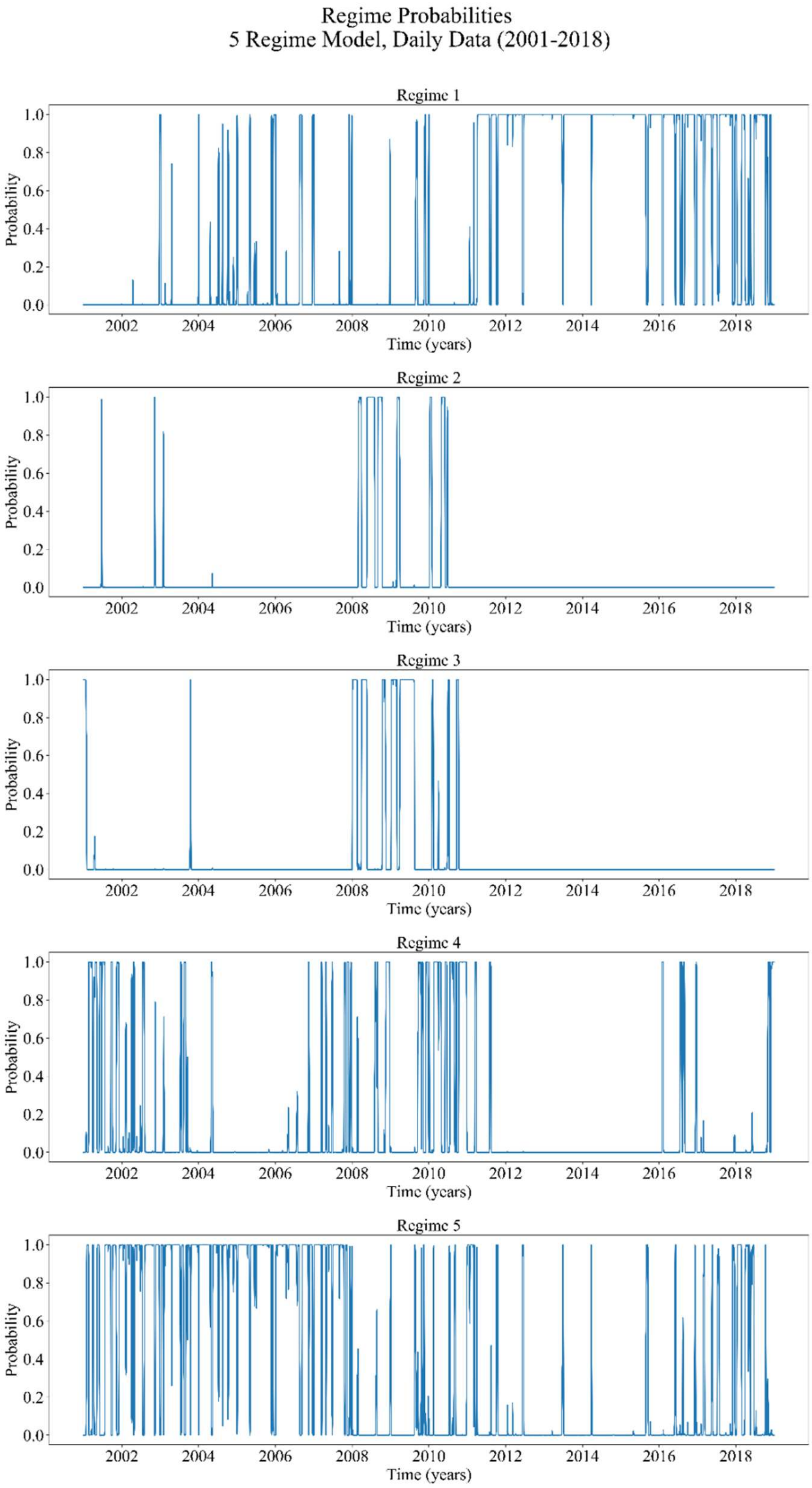
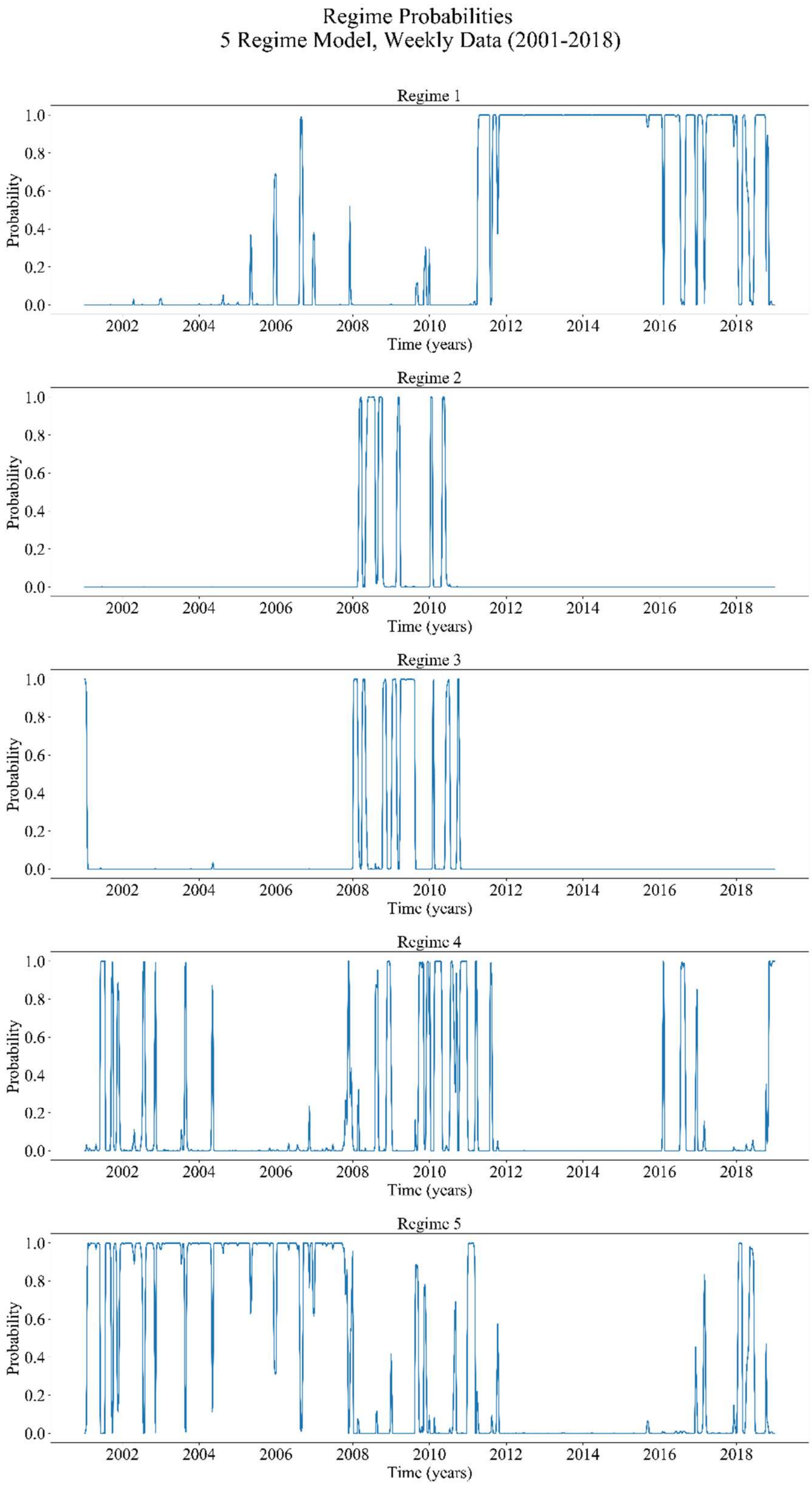


Figure A.1.6: Regime posterior probabilities over time for a 5-regime model estimated using weekly data.



Appendix 2: ICEFR 2019 Conference Paper

This appendix contains a conference paper, presented at the 8th International Conference of Economics and Finance Research, held in Lyon, France from June 18th to 21st 2019. This research paper was the basis for Chapter 2 of the thesis. This paper will also be published in the International Journal of Trade, Economics and Finance.

Trends and Applications of Machine Learning in Quantitative Finance

Sophie Emerson, Ruairi Kennedy, Luke O'Shea, and John O'Brien

Abstract — **Recent advances in machine learning are finding commercial applications across many industries, not least the finance industry. This paper focuses on applications in one of the core functions of finance, the investment process. This includes return forecasting, risk modelling and portfolio construction. The study evaluates the current state of the art through an extensive review of recent literature. Themes and technologies are identified and classified, and the key use cases highlighted. Quantitative investing, traditionally a leading field in adopting new techniques is found to be the most common source of use cases in the emerging literature.**

Index Terms—**Machine Learning, Quantitative Finance, Portfolio Construction, Return Forecasting**

INTRODUCTION

Machine Learning (ML) is a subfield of Artificial Intelligence (AI) that uses statistical techniques that provide computer models with the ability to learn from a dataset, allowing the models to perform specific tasks without explicit programming [1]. ML is being applied to improve function across the finance industry in a wide range of areas including, for example, fraud detection, payment processing and regulation. This research evaluates current and potential applications of machine learning to the investment process. In particular, this includes the development of ML applications for return forecasting, portfolio construction and risk modelling.

The first widespread commercial use cases of artificial intelligence were "expert systems", originating in Stanford in the 1960s [2] and popularised in the 1980s and 1990s. Expert systems were designed to solve complex problems in a specific field, in a manner similar to a subject matter expert. Original expert systems were rule-based programs developed in languages such as LISP and Prolog. In recent years, there has been a significant drop in interest in classic expert systems, as they are superseded by systems incorporating artificial intelligence [3]. AI systems are systems that replicate human thought processes. [4]. Many of these systems are advertised today as cognitive computing systems.

Manuscript received April 19, 2019. This work was supported by State Street Corporation and the authors wish to thank State Street for this support.
Sophie Emerson (sophieemerson@gmail.com), Ruairi Kennedy (r.ocinneide@umail.ucc.ie), and Luke O'Shea (davidluke.oshea@gmail.com) are researchers in the State Street Advanced Technology Centre, Cork University Business School, UCC, Ireland.

John O'Brien (j.obrien@ucc.ie) is a lecturer in the Department of Accounting & Finance, Cork University Business School, UCC, Ireland.

Cognitive computing describes a computer system which mimics human cognitive process in some way, cognitive processes are those that allow individuals to remember, think, learn and adapt [5]. The term has gained recognition in the public domain in recent years, due in large to the introduction of Watson, IBM's cognitive computing system. These systems are constructed by combining computer science with statistical and ML techniques developed over the last century [1]. Watson, in its original form, was a question answering computing system, responding to questions posed in natural language. It was introduced on the television quiz show "Jeopardy!" – where it defeated two of the show's most celebrated contestants in the "IBM Challenge" [6]. Large-scale systems such as Watson combine many techniques [6] to provide "augmented human intelligence" services to users [7]. However, the use of individual techniques, for example deep learning neural networks or reinforcement learning, has found significant success across industry and applications [8-10].

Recently, there has been a proliferation of ML techniques and growing interest in their applications in finance, where they have been applied to sentiment analysis of news, trend analysis, portfolio optimization, risk modelling among many use cases supporting investment management. This paper explores the potential of ML to enhance the investment process. We begin with a broad survey of the area to determine the main programming languages, frameworks and use cases for ML from the perspective of the financial industry. We then focus on ML and its potential applications to quantitative investment. We look at research that has applied ML to the investment process, analysing the technologies used, the functions of the applications, and evidence of potential to improve investment outcomes. Our findings are relevant to both academics and practitioners with interest in investment management, and in particular quantitative investment, by providing a detailed discussion of the latest technologies, their potential uses, and probability of successful application.

The paper is organized as follows. In Section II, we provide an overview of the development of the area as a background for the discussion, this includes the emergence of ML, common algorithms and methodologies, and a review of the evolution and theory of quantitative investing We then describe the research methods in Section III. Section IV provides a detailed description of the current state of the art in the application of ML to investment. We conclude with a discussion of the evidence presented in Section V.

BACKGROUND

Machine Learning

Although variations of ML have long been around, the discipline has developed rapidly in recent years. Many factors have combined to derive this development. Increased computer power has made real time processing feasible for many complex tasks, increased connectivity has driven innovation and automation in the delivery of traditional tasks and services, the potential to extract useful information from the vast amounts of data generated via the internet (Big Data) has led to novel analytical methods. Alongside this, the development of easy to use programming languages, such as Python and R, and ML focused frameworks such as TensorFlow, has contributed to the wide investigation of ML applications in industry. It has already found commercial application across multiple industries from automated trading systems in the finance industry to the health sector where ML algorithms assist decision making in fertility treatments [11]. The success of these applications is driving commercial research into further applications.

Common ML Approaches and Algorithms

Three main approaches to training ML algorithms are recognised; supervised learning, unsupervised learning and reinforcement learning. Supervised learning generates a function that maps inputs to outputs based on a set of training data. The algorithm infers a function linking each set of inputs with the expected, or labelled, output in the training set. Unsupervised learning finds hidden patterns in and draws inferences from unlabelled data. Unsupervised learning provides inputs to models, but does not specify an expected set of outcomes, the outcomes are unlabelled. Reinforcement learning enables algorithms to learn by trial and error, based on feedback from past experiences. Like unsupervised learning, it does not require labelled data. A hybrid system, semi-supervised learning, combines supervised and unsupervised learning, using both labelled and unlabelled data to train models. This is useful where there is limited data or the process of labelling data could introduce biases.

The main research areas in supervised learning are regression and classification (specifying the category or class to which something belongs), this approach is often used in developing predictive models. Regression techniques predict continuous responses using algorithms such as linear regression, decision trees and Artificial Neural Networks (ANNs). Classification techniques predict discrete responses using algorithms such as logistic regression, Support Vector Machines (SVMs) or K-Nearest Neighbors (KNN). The main research area in unsupervised learning is clustering. Clustering refers to grouping objects together, such that objects that are put in the same group are more similar to each other than objects in other groups.

Artificial neural networks have become a key technology in the development of ML. They were first proposed over 75 years ago, inspired by the workings of the human brain [12]. They are a collection of algorithms that replicate the process of a biological brain at the neuron level [1].

There are a number of different classes of artificial neural networks, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and recursive neural networks, among others. CNNs are ideal for things such as image classification and video processing because they're able to identify patterns by focusing on fragments of images. RNNs are better for dealing with things like speech or text analysis because they use time-series information, such as monthly stock price figures to predict next month's figure. GANs have garnered much interest in recent years since they were first introduced in 2014 [13]. GANs are comprised of two neural networks that compete against each other. One neural network generates data similar to the training dataset, and the other tries to evaluate whether data is from the training dataset or generated by the generative network.

Aside from neural networks other well-known ML algorithms include SVMs, KNN and other. SVMs, used for classification and regression analysis, involve finding a hyperplane which minimizes the distance between a set of data points in an n-dimensional space. Bayesian networks are built from probability distributions and use probability laws for prediction and anomaly detection. KNN selects the most similar data points in the training data, this allows the algorithm to classify future data inputs in the same way. Some techniques are better suited to particular tasks than others. This research partly seeks to contribute to this area of knowledge. It is important to evaluate the effectiveness of certain algorithms, to assist in choosing appropriate algorithms for specific tasks in future applications and studies.

The Evolution of Quantitative Investing

Graham and Dodd's *Security Analysis*, published in 1934 following the Wall Street Crash of 1929 is the seminal work on fundamental investing and remains in publication today [14]. It is one of the first books to distinguish investing from speculation, advocating the use of a systematic framework for analysing securities for stock selection.

A systematic approach to portfolio construction and risk analysis was presented in *Portfolio Selection* [15], published in 1952. In this, Markowitz provides a mathematical definition of risk as the standard deviation of return. The approach focused on maximizing portfolio performance by optimizing the trade-off between risk and return. This was the foundation of modern portfolio theory, providing an analytical framework for the construction and analysis of investment portfolios [16], [17].

A quantitative approach to market analysis gained popularity as advances in computing technology made the collection and analysis of large amounts of market data possible. This allowed the development and verification of market models on a scale not previously possible, contributing to significant advances in the understanding of financial markets, including the Capital Asset Pricing Model (CAPM) [18]-[21] and Efficient Market Hypothesis (EMH) [22].

In 1973, Fama and MacBeth used the Center for Research in Security Prices (CRSP) financial dataset (one of the first of its kind) to perform an empirical analysis of the CAPM [23]. They showed that the CAPM provided a good quantitative approximation of the behaviour of security prices while setting a standard for empirical cross-sectional analysis of market data [23].

The empirical support for the EMH, enhanced by the success of market indices, such as the S&P 500, led to the dominant view, particularly in academia, that active investing was futile, as it was impossible to beat a passive investment. In comprehensive literature reviews, [16] and [17] provide evidence that research and empirical evidence that challenged the CAPM and EMH was strongly discouraged. At the same time many examples of research that argued that although difficult, it is possible for active management to beat passive management, by exploiting market inefficiencies not covered by the CAPM and EMH. Strategies based on risk factor models, first explored by Rosenberg [24] and Ross [25] in the 1970s, surged in popularity [26] after the publication of the Fama-French three-factor model [27].

From Markowitz portfolio optimization to CAPM, EMH and factor models more recently, quantitative investors have shown that they are willing to embrace new techniques and strategies. A key argument for applying ML techniques to financial problems is that ML methods capture non-linear relationships [28] in the data. Non-linear methods are required to model data where outputs are not directly proportional to the inputs [29] and many traditional analysis methods assume a linear relationship, or a non-linear model that can be simplified to a linear model. Typical examples of well-established non-linear ML methods include KNN, and ANN [20].

ML has been applied with positive results across many areas of quantitative investing, including portfolio optimization [30], [31], factor investing [32], bond risk predictability [29], derivative pricing, hedging and fitting [33], and back-testing [34]. The results section contains a comprehensive summary of papers where ML techniques are applied to areas of quantitative finance.

METHODOLOGY

Initially, a broad search was conducted to identify the major themes related to ML. This search yielded information on the popular use cases and technologies. This information informed a second, more focused investigation of relevant material. Here, the aim was to draw connections between popular use cases in finance and current ML techniques.

As quality and scope of published research can vary widely, measures were taken to reduce the possibility of including unreliable information in the final dataset. Before inclusion in the concept matrix, each paper was assessed on quality. This was achieved by using a variety of quality indicators including; the citation count, the quality of an institute's research activities associated with the paper, bias created from funding sources, and the impact factor of the journal.

An appropriate search strategy was devised and carried out based on the main topics that were identified during the first investigation of the literature. The arXiv and SSRN databases were searched to ensure that the most up-to-date research papers were included. However, as these are not peer-reviewed papers, extra care was taken to ensure that the papers were from reputable authors, focusing on the quality of authors' previous publications. The topic phrases used in search were "portfolio management", "stock market forecasting", and "risk management". All of these topic phrases were used in conjunction with the key phrase "machine learning" in an attempt to return only relevant research papers. The purpose of searching by topic was to identify which technologies are widely and effectively used within each area. As we are evaluating the current state of the art, we wanted to ensure that only recent papers were included. Thus, we only included papers that were submitted in 2015 or later. From the initial search we collected a total of 118 papers. After an initial review of abstracts, papers that were not relevant to machine learning in finance (specifically investing) were removed. Any papers that were duplicates under more than one search topic were kept under the topic that appeared most relevant. Papers were then assessed in relation to their quality using the quality indicators mentioned above. This reduced the number of papers to 55.

RESULTS

Popular Machine Learning Use Cases and Algorithms

A concept-centric matrix was utilised initially to identify which areas commonly use machine learning techniques. Recurring concepts and themes were noted and counted across a sample of 67 papers identified. An initial list of recurring themes was identified and analysed. Some themes, such as 'Geopolitics' were removed as they were deemed irrelevant due to the lack of research on the topic. A list of the most recurring themes with relevance to ML is presented in Table I.

TABLE I: RECURRING THEMES FROM THE LITERATURE REVIEW.

Theme	**References**
Return Forecasting	21
Portfolio Construction	12
Ethics	8
Fraud Detection	8
Decision Making	8
Language Processing	7
Sentiment Analysis	7

The most common use-cases identified were return forecasting and portfolio construction. Quantitative methods were introduced to finance through the equity market and it is unsurprising that it should lead the way in incorporating the latest advances in its processes. A large number of the papers above also discussed risk modelling. This led us to take return forecasting, portfolio construction, and risk modelling as our three core topics. The most popular ML techniques identified in the papers researched are presented in Table II, as well as a breakdown of the different acronyms used in the table.

TABLE II: POPULAR TECHNIQUES FEATURED IN MACHINE LEARNING AND FINANCE PAPERS

	MLP	SVM	LSTM	GRU	RNN	CNN	RF	GPR	LR
Return Forecasting	7	5	4	2	-	1	2	-	-
Portfolio Construction	7	2	3	1	1	1	4	2	1
Risk Modelling	6	2	2	1	1	1	4	3	4

MLP	Multilayer Perceptron
SVM	Support Vector Machine
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
RNN	Recurrent Neural Network (basic)
CNN	Convolutional Neural Network
RF	Random Forests/Decision Trees
GPR	Gaussian Process Regression
LR	Logistic Regression

Many techniques used in the papers only appear once, some twice. Since the purpose of this paper is to identify the most popular machine learning techniques used in finance, specifically in the topics above, only techniques which appeared in at least three papers were included in Table II. We also decided to include RNN, although it is only mentioned explicitly in two papers, it appears implicitly more frequently as both LSTM and GRU are subsets of the technology.

Artificial neural networks are used in all three areas of finance studied, with a standard feedforward network (MLP) being the most common. Useful results are found from networks that range from small to very large networks (deep neural networks). There is also evidence of preferences for some techniques in particular areas. For example, Gaussian process regression is used in both portfolio construction and risk modelling but has not been applied to return forecasting.

Summary of Key Insights from Recent Papers

The paper selection included ML papers published in recent years as well as papers yet to be published by established authors from reputable institutions. These papers have been submitted for publication and are awaiting acceptance. The most recent studies in this field were included to help evaluate the cutting edge and state of the art of the use of ML for financial applications.

I. *Portfolio Construction*

Portfolio construction is the process of combining return forecasts and risk models to create an optimum portfolio given an investor's constraints. A variety of ANN methodologies are applied to the portfolio optimisation problem, often outperforming traditional optimisation techniques. Deep learning reappeared a number of times during this search in the context of portfolio construction. Deep learning refers to models that consist of multiple layers or stages of nonlinear information processing (for example, a neural network with many hidden layers) [35]. Both hierarchical clustering and reinforcement learning were used to improve portfolio diversification. Multiple papers discuss the method of applying Markov models to predict the performance of stocks. Markov models are a type of ML method that model variables that change randomly through time. The complicated nature of the global market makes using this type of model a viable option.

- The authors present a deep learning framework for portfolio design, applying their framework to the stocks in the IBB index, demonstrating that their portfolio weighted using deep learning outperformed the index [31].
- The author outlines a reinforcement learning solution for a rational risk-averse investor seeking to maximize expected utility of final wealth, giving an example of a Q-learning agent exploiting an approximate arbitrage in a simulation [36].
- The authors of both papers make use of hierarchical clustering algorithms for constructing diversified portfolios. The portfolios are constructed using variations of risk parity [30] and equal risk contribution methods [37] which take the hierarchical correlation structure of the assets into account. The portfolios constructed are shown to have superior diversification and out-of-sample risk adjusted performance.
- The authors make use of convex analysis techniques to devise an optimal portfolio coupled with a Hidden Markov Model (HMM) used to estimate growth rates in the market model, which achieves improved results over a simple model using geometric Brownian motions [38].
- The authors provide an overview of the financial applications of Gaussian processes and Bayesian optimisation, providing examples for forecasting the yield curve with Gaussian processes, and using Bayesian optimisation to build an online trend-following portfolio optimisation strategy [39].
- The authors compare the use of Feature Salient Hidden Markov Models (FSHMM) and HMM for constructing factor investing portfolios. The FSHMM selects relevant factors for use from a pool of available factors, while the HMM uses the whole pool of factors. Both models outperformed benchmark portfolios, with the FSHMM portfolio showing better performance [40]
- The authors use factors as inputs to deep neural network, SVM and random forest models for predicting stock returns. While their research shows the effectiveness of a deep learning model, more significantly they used Layer-wise Relevance Propagation (LRP) to determine individual factor contributions to the neural network's prediction [41].
- The authors create a non-linear multi-factor model using LSTM to estimate the non-linear function. As in the previous paper the authors make use of LRP to identify which factors contribute to the model. The performance of the LSTM model is compared to the neural network model used in [32] and gives superior returns [42].
- The authors examine the use of three deep reinforcement learning algorithms, Deep Deterministic Policy Gradient (DDPG), Proximal Policy Optimization (PPO) and Policy Gradient (PG), in managing a portfolio of assets in the Chinese stock market. They determine that training conditions used in game playing and robot control are unsuitable for use with portfolio management, finding that DDPG and PPO gave unsatisfying performance in the training process. They propose the use of adversarial training methods and employ a revised PG algorithm which outperforms a Uniform Constant Rebalanced Portfolio (UCRP) benchmark [43].
- The authors employ models constructed using Gaussian processes and Monte Carlo Markov Chains which learn optimal strategies from historical data, based on user-specified performance metrics (e.g. excess return to the market index, Sharpe ratio, etc.). This approach addresses the inverse problem of Stochastic Portfolio Theory – devising suitable investment strategies that meet the desired investment objective, when initially

given a user-defined portfolio selection. The models outperform the benchmark in-sample and out-of-sample for absolute terms (returns) and also after adjusting for risk (Sharpe ratio) [44].

- The author provides an ML framework for estimating optimal portfolio weights. They apply this framework using three ML methods – Ridge and Lasso regression, and two newly introduced methods; Principal Component regression, Spike and Slab regression. All methods outperform the mean-variance, minimum-variance, and equal weight portfolios. [45].
- The authors propose a way to find the risk budgeting portfolio by using optimisation algorithms to find a solution to the logarithmic barrier problem. They use algorithms such as cyclical coordinate descent, alternating direction method of multipliers (ADMM), proximal operators, and Dykstra's algorithm [46].
- The authors present a financial-model-free reinforcement learning framework as a solution to the portfolio management problem. The study tests the proposed framework with the following neural networks: CNN, a basic RNN, and LSTM [47].

II. *Return Forecasting*

Return forecasting, predicting the investment return from an asset or asset class, is central to investment management and features highly in the literature. Many types of ANN are tested on their ability to forecast returns. Deep neural networks, CNNs, LSTMs are all applied to the problem of return forecasting. In one theme, the new ML technology is applied to improve forecasts made from traditional inputs, such as fundamental accounting data or technical indicators. A second approach uses ML to extract new inputs from alternative data, such as sentiment from news data. Finally, authors predict movement at market level rather than at the level of individual securities, for example using ML to identify states.

- The authors use a CNN strategy to analyse and detect price movement patterns in high-frequency limit order book data. Multilayer neural network methods and SVMs were also considered. However, they conclude the CNNs provide better performance for this task [48].
- The authors implement several ML algorithms to predict future price movements using limit order book data. They employ two feature learning methods: Autoencoders, and Bag of Features. They compare three different classifiers: SVM, a Single Hidden Layer Feedforward Neural Network (SLFNN), and an MLP. They test the performance of the classifiers with an anchored walk forward analysis, to determine if the models can capture temporal information, as well as a hold-out per stock method, to determine if the models can learn features that can be applied to previously unseen stocks. The results from the MLP are better than the other classifiers. However, the use of the Autoencoder and Bag of Features in combination with the MLP lead to fewer correct predictions [49].
- The authors introduce a novel Temporal Logistic Neural Bag-of-Features approach, that can be used to tackle the challenges that come with data of a high dimensionality, in this case high-frequency limit order book data [50].
- The authors train a deep neural network on reported fundamental data from publicly traded companies (revenue, operating income, debt etc.). The model forecasts future fundamental data based on a trailing 5-years window. A value investing factor strategy based on forecasted fundamental data outperforms a traditional value factor investing

approach with a compounded annual return of 17.1% vs 14.4% for a standard factor model [51].

- The authors create a simple buy-hold-sell strategy to predict direction of movement for 43 CME listed commodities and FX futures based on an ANN trained on a multitude of features for each instrument designed to capture co-movements and historical memory in the data. An average prediction accuracy of 42% is achieved across all instruments, with higher accuracies achieved for certain instruments [52].
- The authors use a random forest model to predict the direction of stock prices based on price information and a number of momentum indicators (Relative Strength Index, Moving Average Convergence Divergence, Stochastic Oscillator, Williams %R, On Balance Volume, and Price Rate of Change). The algorithm is shown to outperform existing algorithms found in the literature [53].
- The authors provide a sentiment analysis dictionary which they use to predict stock movements in the pharmaceutical market sector. With this model they achieve an accuracy of 70.59%. [54]
- The authors present a methodology to define, identify, classify and forecast market states. They use a Triangulated Maximally Filtered Graph network to filter information, and simple logistic regression for predicting market states. They compare five models, with a Gaussian Mixture Model as their baseline. All five models outperform the baseline in terms of risk/return significance [55].
- The authors compare five ANN models for forecasting stock prices: a standard neural network using back propagation, a Radial Basis Function (RBF), a General Regression Neural Network (GRNN), SVM Regression (SVMR), and Least Squares SVM Regression (LS-SVMR). However, they compare the models on just three stocks: Bank of China, Vanke A, and Kweichou Moutai. The standard neural network using back propagation outperforms all of the other models across all three stocks, in terms of both Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). [56]
- The authors use 25 risk factors as inputs to ML stock returns prediction models. Results show that deep neural networks generally outperform shallow neural networks, and the best networks also outperform representative machine learning models [57].
- The author employs ANNs to predict product demand for weather sensitive products in Walmart stores around the time of major weather events [58].
- The authors implement a Gaussian Naïve Bayes Classifier for prediction based on sentiment analysis of Twitter data. The data used was obtained from Twitter and pertained to the 2014 FIFA world cup. Their framework obtained an accuracy and Area Under the curve of the Receiver Operating Characteristic (AUROC) of around 80% and an 8% marginal profit when tested [59].

III. Risk

Three different themes are identified under the broad heading of risk. The first attempts to employ ML to improve traditional measures of risk used in the mean variance framework. The second theme looks for companies at risk of default or bankruptcy. Techniques such as natural language processing are used to identify words that indicate higher risk. The final theme uses ML to develop hedging strategies. Some authors look at identifying what selection of ML methods is best for risk modelling problems.

- The authors use k-means clustering to construct risk models by clustering stock returns normalized by standard deviation squared and adjusted by mean absolute deviation using a method proposed in [60]. They demonstrate that this ML approach outperforms statistical risk models [61] in quantitative trading applications [62].

- The authors present a framework for hedging a portfolio of derivatives in the presence of market frictions such as transaction costs, market impact, liquidity constraints or risk limits [63].
- The authors show how Gaussian Process Regression can assist in pricing and hedging a Guaranteed Minimum Withdrawal Benefit (GMWB) Variable Annuity with stochastic volatility and stochastic interest rate [64].
- The authors show that machine learning can be as effective as other existing algorithms at solving difficult hedging problems in moderate dimension. They use techniques such as a modified LSTM neural network to calculate their hedging strategies [65].
- The authors aim to explore the optimal model for business risk prediction. They attempt to do this using XGBoost, and by simultaneously examining feature selection methods and hyper-parameter optimization in the modelling procedure [66].
- The authors try to predict daily stock volatility using news and price data. Their model, which utilizes a Bidirectional Long Short-Term Memory (BiLSTM) neural network and stacked LSTM's, outperforms the well-known Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model in all sectors analysed (financial, health care, etc.) [67].
- The authors exploit a heterogeneous information network of 35,657 global firms to improve the predictive performance for firms likely to be added to a blacklist. Blacklists are used to keep track of entities that have unacceptable problems, such as financial or environmental issues. Blacklists help keep portfolios profitable and "green". Their model consists of a simple MLP with thirty hidden units [68].
- The authors estimate corporate credibility of Chinese companies using a CNN and natural language processing. They use Latent Dirichlet Allocation to summarise the text of news articles and use a CNN to extract the most important words from each topic. The CNN learns how news articles may reflect the credibility of a company though the wording of articles and word occurrence. They verify their model works by building a negative rating system and showing a correlation between their model's results and the negative rating [69].
- The authors compare different strategies for solving a variation of the multi-armed bandit problem. In their version of the problem, the learner can pull several arms simultaneously, or none at all. This could easily be applied to assist in investment decisions. Out of the strategies compared, Bayes-UCB-4P and TS-4P perform the best [70].
- The authors compare several ML algorithms: Logistic Regression, K-Dimensional Tree (K-D Tree), SVM, Decision Trees, AdaBoost, ANN, and Gaussian Processes (GP) for forecasting business failures (corporate bankruptcy). Models are compared on datasets of manufacturing companies in Korea and Poland. All of the models are compared on their performance when combined with different dimensionality reduction techniques. The techniques used are: Principal Component Analysis (PCA), Linear Discriminate Analysis (LDA), Isometric Feature Mapping (ISOMAP), and Kernel PCA. On the Korean dataset, all models perform similarly. K-D Tree, SVM, and GP perform best over all of the dimensionality reduction methods used. On the Polish dataset, the linear regression model performs the best. Although having a lower accuracy than some of the other models, it is the best performing method when compared over other results such as

precision, recall, F1 score, and AUC (Area Under Curve) [71].

DISCUSSION

Strategy Development & Analysis

The results of the literature search demonstrate that there is a wide range of ML techniques being successfully applied to many areas in the development of quantitative investing strategies, outperforming traditional benchmarks, previously used techniques and algorithms in many cases. Algorithms that assume a linear relationship between data can result in reduced accuracy. [28] highlights this issue in terms of many of the econometric models employed by finance academics and investment managers. The author argues for the use of more advanced mathematical models and ML techniques such as unsupervised learning that are capable of modelling complex non-linear relationships in financial systems.

Taking factor investing as an example of this, [72] and [73] make use of statistical algorithms to show that many factors discovered over the last number of years (particularly those found using empirical evidence) can be considered inaccurate or invalid. In the aptly named paper, *Taming the Factor Zoo*, a double selection LASSO ML method was used to analyse the contribution and usefulness of individual factors amongst the large number available today [74]. LASSO (Least Absolute Shrinkage and Selection Operator) is a regression analysis method capable of reducing the dimensionality of a large sample while selecting variables significant to the final result [75]. In [57] the author uses twenty-five factors as model inputs, comparing the use of shallow and deep neural networks, as well as SVMs and random forests for predicting stock returns, finding the deep neural networks (more layers) superior to the other methods. Using a similar approach [41] uses factors as inputs to deep neural network, SVM and random forest models for predicting stock returns. While their research again showed the effectiveness of a deep learning model, more significantly they used layer-wise relevance propagation to determine individual factors contributions to the neural network's prediction. In these cases, not only has ML been used to develop investment strategies, but also to detect which input features were significant and which were not.

The use of Alternative Data

The use of ML for the analysis and application of alternative data for example, sentiment analysis, supply chain data etc. has opened up opportunities for new investment strategies. As seen in Table I, sentiment analysis was identified as a popular use case for ML. [17] provides a thorough overview of the growth of big data and sentiment analysis research over the last 30 years, highlighting the use of techniques such as NLP, SVMs and ANNs for the analysis of news, conference calls, reports, and social media activity. They concluded that to date, sentiment information has provided short-term, easy to exploit insights but long-term persistent insights are hard to achieve (falling in line with EMH). [16] acknowledges the effectiveness of big data for the modern fundamental investor, as it can provide insights and improve decision making by widening their research capabilities. This sentiment is echoed in [28] where the author makes reference to the recently emerged term "quantamental" – describing a fundamentally leaning investor who manages their portfolio based on data-driven insights provided by ML algorithms. Examples of ML and alternative data being applied together in the results section mainly fall under return forecasting or risk modelling, where decisions may be made based on good or bad news [54], weather [58], or social media sentiment [59].

Choosing Machine Learning Algorithms

It is important to understand the relevant factors that contribute to the choice of ML algorithms, given the wide range available. These factors include accuracy, training time, linearity, number of parameters, the number of features and the structure of the data [76]. Some systems do not need a high level of accuracy. Estimates may be sufficient, for example, when calculating different route times for a journey. Model training times can also vary hugely between algorithms, making some algorithms more appealing than others when under time constraints. Many algorithms assume a linear relationship between input and output (linear regression, logistic regression, SVMs). This can result in reduced accuracy when dealing with non-linear problems. The number of parameters an algorithm has can indicate its flexibility, but also indicates that more time and effort may be required to find optimal values for training the model. The number of features can also be overwhelming for some algorithms. This is particularly a problem with textual data, where the number of words in the dictionary vastly outweighs the number of words in say, a paragraph being used for sentiment analysis. It's important to consider the structure of the data and the specific problem, as some algorithms are better suited for certain problems and data structures [77].

Backtesting & Strategy Verification

While ML techniques can provide superior performance, financial data is notorious for having a low signal-to-noise ratio, which can lead to the detection of false patterns and results. Backtesting protocols have been proposed to tackle this [78]. ML solutions have also been applied to this problem. In [34] the authors present an unsupervised learning strategy which makes use of a modified k-means clustering algorithm to extract the number of uncorrelated trials from a series of backtests, which can be used in estimating the probability of false positives and estimating the expected value of the maximum Sharpe ratio. While in [79] the authors use a machine learning strategy for backtesting and the evaluation of automated trading strategies which is trained on a number of performance and risk metrics, demonstrating that this strategy outperforms standard metrics such as Sharpe ratio out-of-sample.

The development of new backtesting strategies and protocols is welcome and necessary, especially taking into account recent "black box" criticisms by leading deep learning researchers regarding a lack of testing and reproducibility in the field of ML. In their acceptance speech after winning the "test-of-time" award at NIPS, the leading AI conference, the authors of [80] compared much of recent ML research to "alchemy", highlighting a situation where algorithms were being created and trained using trial and error methods, with the researchers unable to explain the fundamental operation. They later published a paper highlighting instances of this [81].

CONCLUSION

As the previous section discusses, ML offers an opportunity for more complex financial analysis than was previously possible. The literature shows that quantitative investors have embraced new tools and techniques as they have emerged [16], [17].

There is a growing body of literature applying ML techniques to investment problems. Varieties of ML methods have been applied to areas of quantitative finance– the most popular methods are MLPs, followed by SVMs, and LSTM. ML has been applied to problems in areas such as return forecasting, portfolio construction, and risk modelling.
These ML methods utilize traditional financial data, as well as making use of new types of alternative data. Big data is providing new datasets that need to be analysed and ML techniques are capable of modelling complex (non-linear) relationships and analysing new data.

[28] notes the recent trend of traditional hedge funds hiring an increasing proportion of STEM graduates for portfolio construction positions, as they possess the required mathematical skillset for performing complex analysis and computer modelling. An understanding of machine learning, as well as the languages (Python, R, etc.) and frameworks (e.g. TensorFlow) needed to construct complex models could certainly be considered advantageous for any quantitative investor looking for an edge.

REFERENCES

[1] P. Domingos, "A few useful things to know about machine learning," *Communications of the ACM,* vol. 55, no. 10, pp. 78-87, 2012.

[2] R. K. Lindsay, B. G. Buchanan, E. A. Feigenbaum, and J. Lederberg, "DENDRAL: a case study of the first expert system for scientific hypothesis formation," *Artificial intelligence,* vol. 61, no. 2, pp. 209-261, 1993.

[3] W. P. Wagner, "Trends in expert system development: A longitudinal content analysis of over thirty years of expert system case studies," *Expert systems with applications,* vol. 76, pp. 85-96, 2017.

[4] S. Russell and P. Norvig, *Artificial Intelligence A Modern Approach Third Edition*. Prentice Hall, 2009.

[5] D. Modha and C. Witchalls, "A computer that thinks," vol. 224, ed: New Scientist Ltd., 2014, pp. 28-29.

[6] D. Ferrucci *et al.*, "Building Watson: An overview of the DeepQA project," *AI magazine,* vol. 31, no. 3, pp. 59-79, 2010.

[7] H. Reynolds, "AI? Or cognitive computing?," *KM World,* Article vol. 26, no. 9, pp. 4-5, 2017.

[8] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing,* vol. 234, pp. 11-26, 2017/04/19/ 2017.

[9] R. Rana and F. S. Oliveira, "Dynamic pricing policies for interdependent perishable products or services using reinforcement learning," *Expert Systems with Applications,* vol. 42, no. 1, pp. 426-436, 2015.

[10] G. Choy *et al.*, "Current applications and future impact of machine learning in radiology," *Radiology,* vol. 288, no. 2, pp. 318-328, 2018.

[11] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

[12] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics,* vol. 5, no. 4, pp. 115-133, 1943.

[13] I. Goodfellow *et al.*, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672-2680.

[14] B. Graham and D. Dodd, *Security Analysis: Foreword by Warren Buffett*. McGraw-Hill Professional, 2008.

[15] H. Markowitz, "Portfolio selection," *The journal of finance,* vol. 7, no. 1, pp. 77-91, 1952.

[16] R. N. Kahn, *The Future of Investment Management*. CFA Institute Research Foundation, 2018.

[17] Y. L. Becker and M. R. Reinganum, *The Current State of Quantitative Equity Investing*. CFA Institute Research Foundation, 2018.

[18] W. F. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *The journal of finance,* vol. 19, no. 3, pp. 425-442, 1964.

[19] J. Mossin, "Equilibrium in a capital asset market," *Econometrica: Journal of the econometric society,* pp. 768-783, 1966.

[20] J. Lintner, "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," in *Stochastic Optimization Models in Finance*: Elsevier, 1975, pp. 131-155.

[21] C. W. French, "The Treynor capital asset pricing model," *Journal of Investment Management,* vol. 1, no. 2, pp. 60-72, 2003.

[22] B. G. Malkiel and E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *The journal of Finance,* vol. 25, no. 2, pp. 383-417, 1970.

[23] E. F. Fama and J. D. MacBeth, "Risk, return, and equilibrium: Empirical tests," *Journal of political economy,* vol. 81, no. 3, pp. 607-636, 1973.

[24] B. Rosenberg, "Extra-market components of covariance in security returns," *Journal of Financial and Quantitative Analysis,* vol. 9, no. 2, pp. 263-274, 1974.

[25] S. A. Ross, "The arbitrage theory of capital asset pricing," in *HANDBOOK OF THE FUNDAMENTALS OF FINANCIAL DECISION MAKING: Part I*: World Scientific, 2013, pp. 11-30.

[26] J. H. Cochrane, "Presidential Address: Discount Rates," *The Journal of Finance,* vol. 66, no. 4, pp. 1047-1108, 2011.

[27] E. F. Fama and K. R. French, "Common risk factors in the returns on stocks and bonds," *Journal of financial economics,* vol. 33, no. 1, pp. 3-56, 1993.

[28] M. Lopez de Prado, "Mathematics and economics: a reality check," 2016.

[29] D. Bianchi, M. Büchner, and A. Tamoni, "Bond risk premia with machine learning," *Available at SSRN 3232721,* 2018.

[30] M. Lopez de Prado, "Building diversified portfolios that outperform out-of-sample," *Journal of Portfolio Management,* 2016.

[31] J. Heaton, N. Polson, and J. H. Witte, "Deep learning for finance: deep portfolios," *Applied Stochastic Models in Business and Industry,* vol. 33, no. 1, pp. 3-12, 2017.

[32] K. Nakagawa, T. Uchida, and T. Aoshima, "Deep Factor Model," *arXiv preprint arXiv:1810.01278,* 2018.

[33] J. De Spiegeleer, D. B. Madan, S. Reyners, and W. Schoutens, "Machine learning for quantitative finance: fast derivative pricing, hedging and fitting," *Quantitative Finance,* pp. 1-9, 2018.

[34] M. Lopez de Prado and M. J. Lewis, "Detection of False Investment Strategies Using Unsupervised Learning Methods," *Available at SSRN: https://ssrn.com/abstract=3167017 or http://dx.doi.org/10.2139/ssrn.3167017,* 2018.

[35] L. Deng and D. Yu, "Deep learning: methods and applications," *Foundations and Trends® in Signal Processing,* vol. 7, no. 3–4, pp. 197-387, 2014.

[36] G. Ritter, "Machine learning for trading," 2017.

[37] T. Raffinot, "Hierarchical clustering based asset allocation," *Available at SSRN 2840729,* 2017.

[38] A. Al-Aradi and S. Jaimungal, "Active and Passive Portfolio Management with Latent Factors," *arXiv preprint arXiv:1903.06928,* 2019.

[39] J. Gonzalvez, E. Lezmi, T. Roncalli, and J. Xu, "Financial Applications of Gaussian Processes and Bayesian Optimization," *arXiv preprint arXiv:1903.04841,* 2019.

[40] E. Fons, P. Dawson, J. Yau, X.-j. Zeng, and J. Keane, "A novel dynamic asset allocation

system using Feature Saliency Hidden Markov models for smart beta investing," *arXiv preprint arXiv:1902.10849,* 2019.

[41] K. Nakagawa, T. Uchida, and T. Aoshima, "Deep factor model," in *ECML PKDD 2018 Workshops*, 2018, pp. 37-50: Springer.

[42] K. Nakagawa, T. Ito, M. Abe, and K. Izumi, "Deep Recurrent Factor Model: Interpretable Non-Linear and Time-Varying Multi-Factor Model," in *AAAI-19 Workshop on Network Interpretability for Deep Learning*, Honolulu, Hawaii, USA, 2019: arXiv preprint arXiv:1901.11493.

[43] Z. Liang, K. Jiang, H. Chen, J. Zhu, and Y. Li, "Deep Reinforcement Learning in Portfolio Management," *arXiv preprint arXiv:1808.09940,* 2018.

[44] Y.-L. K. Samo and A. Vervuurt, "Stochastic Portfolio Theory: a machine learning perspective," presented at the Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence, Jersey City, New Jersey, USA, 2016.

[45] D. Kinn, "Reducing Estimation Risk in Mean-Variance Portfolios with Machine Learning," *arXiv preprint arXiv:1804.01764,* 2018.

[46] J.-C. Richard and T. Roncalli, "Constrained Risk Budgeting Portfolios: Theory, Algorithms, Applications & Puzzles," *arXiv preprint arXiv:1902.05710,* 2019.

[47] Z. Jiang, D. Xu, and J. Liang, "A deep reinforcement learning framework for the financial portfolio management problem," *arXiv preprint arXiv:1706.10059,* 2017.

[48] A. Tsantekidis, N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks," in *2017 IEEE 19th Conference on Business Informatics (CBI)*, 2017, vol. 01, pp. 7-12.

[49] P. Nousi *et al.*, "Machine learning for forecasting mid price movement using limit order book data," *arXiv preprint arXiv:1809.07861,* 2018.

[50] N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Temporal Logistic Neural Bag-of-Features for Financial Time series Forecasting leveraging Limit Order Book Data," *arXiv preprint arXiv:1901.08280,* 2019.

[51] J. Alberg and Z. C. Lipton, "Improving Factor-Based Quantitative Investing by Forecasting Company Fundamentals," *arXiv preprint arXiv:1711.04837,* 2017.

[52] M. Dixon, D. Klabjan, and J. H. Bang, "Classification-based financial markets prediction using deep neural networks," *Algorithmic Finance,* no. Preprint, pp. 1-11, 2017.

[53] L. Khaidem, S. Saha, and S. R. Dey, "Predicting the direction of stock market prices using random forest," *arXiv preprint arXiv:1605.00003,* 2016.

[54] D. Shah, H. Isah, and F. Zulkernine, "Predicting the Effects of News Sentiments on the Stock Market," in *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 4705-4708: IEEE.

[55] P. F. Procacci and T. Aste, "Forecasting market states," *Available at SSRN: https://ssrn.com/abstract=3215945 or http://dx.doi.org/10.2139/ssrn.3215945,* 2018.

[56] Y.-G. Song, Y.-L. Zhou, and R.-J. Han, "Neural networks for stock price prediction," *arXiv preprint arXiv:1805.11317,* 2018.

[57] M. Abe and H. Nakayama, "Deep Learning for Forecasting Stock Returns in the Cross-Section," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2018, pp. 273-284: Springer.

[58] E. Taghizadeh, "Utilizing artificial neural networks to predict demand for weather-sensitive products at retail stores," *arXiv preprint arXiv:1711.08325,* 2017.

[59] L. Le, E. Ferrara, and A. Flammini, "On predictability of rare events leveraging social media: a machine learning perspective," in *Proceedings of the 2015 ACM on Conference on Online Social Networks*, 2015, pp. 3-13: ACM.

[60] Z. Kakushadze and W. Yu, "Statistical industry classification," *Journal of Risk & Control,* vol. 3, no. 1, pp. 17-65, 2016.

[61] Z. Kakushadze and W. Yu, "Statistical risk models," *The Journal of Investment Strategies,* vol. 6, no. 2, pp. 1-40, 2017.

[62] Z. Kakushadze and W. Yu, "Machine Learning Risk Models," *Journal of Risk & Control,* vol. 6, no. 1, pp. 37-64, 2019.

[63] H. Buehler, L. Gonon, J. Teichmann, and B. Wood, "Deep hedging," *Quantitative Finance,* pp. 1-21, 2019.

[64] L. Goudenège, A. Molent, and A. Zanette, "Gaussian Process Regression for Pricing Variable Annuities with Stochastic Volatility and Interest Rate," *arXiv preprint arXiv:1903.00369,* 2019.

[65] S. Fecamp, J. Mikael, and X. Warin, "Risk management with machine-learning-based algorithms," *arXiv preprint arXiv:1902.05287,* 2019.

[66] Y. Wang and X. S. Ni, "A XGBoost risk model via feature selection and Bayesian hyper-parameter optimization," *International Journal of Database Management Systems (IJDMS) preprint arXiv:1901.08433,* 2019.

[67] M. Sardelich and S. Manandhar, "Multimodal deep learning for short-term stock volatility prediction," *arXiv preprint arXiv:1812.10479,* 2018.

[68] R. Hisano, D. Sornette, and T. Mizuno, "Predicting Adverse Media Risk using a Heterogeneous Information Network," *arXiv preprint arXiv:1811.12166,* 2018.

[69] M. Zhang, Z. Luo, and H. Lu, "Latent Dirichlet Allocation with Residual Convolutional Neural Network Applied in Evaluating Credibility of Chinese Listed Companies," *arXiv preprint arXiv:1811.11017,* 2018.

[70] M. Achab, S. Clémençon, and A. Garivier, "Profitable Bandits," presented at the Proceedings of The 10th Asian Conference on Machine Learning, Proceedings of Machine Learning Research, 2018. Available: http://proceedings.mlr.press

[71] J. C. Chow, "Analysis of Financial Credit Risk Using Machine Learning," *arXiv preprint arXiv:1802.05326,* 2018.

[72] C. Harvey and Y. Liu, "Lucky factors," 2017.

[73] C. R. Harvey, Y. Liu, and H. Zhu, "… and the Cross-Section of Expected Returns," *The Review of Financial Studies,* vol. 29, no. 1, pp. 5-68, 2016.

[74] G. Feng, S. Giglio, and D. Xiu, "Taming the factor zoo," 2017.

[75] A. Belloni, V. Chernozhukov, and C. Hansen, "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies,* vol. 81, no. 2, pp. 608-650, 2014.

[76] R. Barga, V. Fontama, and W. H. Tok, "Introducing Microsoft Azure Machine Learning," in *Predictive Analytics with Microsoft Azure Machine Learning*: Springer, 2015, pp. 21-43.

[77] P. Harrington, "Machine learning in action," *Shelter Island, NY: Manning Publications Co,* 2012.

[78] R. Arnott, C. R. Harvey, and H. Markowitz, "A Backtesting Protocol in the Era of Machine Learning," *The Journal of Financial Data Science,* vol. 1, no. 1, pp. 64-74, 2019.

[79] T. Wiecki, A. Campbell, J. Lent, and J. Stauth, "All That Glitters Is Not Gold: Comparing Backtest and Out-of-Sample Performance on a Large Cohort of Trading Algorithms," *The Journal of Investing,* vol. 25, no. 3, pp. 69-80, 2016.

[80] B. Recht and A. Rahimi, "Reflections on random kitchen sinks, 2017," ed, 2017.

[81] D. Sculley, J. Snoek, A. Wiltschko, and A. Rahimi, "Winner's Curse? On Pace, Progress, and Empirical Rigor," 2018.

Appendix 3: CFA European Quant Awards 2019 - 2nd Place Finish

This appendix contains a research report which achieved 2nd place at the CFA European Quant Awards 2019. The Quant Awards are a competition where students and interns are invited to submit an original research report in quantitative finance focusing on a chosen issue in portfolio management. The report should focus on the importance and practical application of the research rather than the techniques used. It must be original work carried out by the candidate, and can include an end of studies dissertation, internship report or work carried out specifically for the Award. The report submitted was based on the contents of Chapters 3-5 of the thesis.

Dynamic Regime-Based Asset Allocation using International Equity Flow Data

Ruairi Kennedy
State Street Advanced Technology Centre, University College Cork
r.ocinneide@umail.ucc.ie

Abstract

Regime switching models have been shown to capture changes of behaviour in financial markets. Often these changes of behaviour persist for some time, and the underlying regimes associated the changes can be used to capture dynamics of asset returns. Similarly, international portfolio flows have been shown to be persistent and to have a relationship with equity returns.

The purpose of this paper is to identify investment regimes based on cross-border and regional equity flow data using hidden Markov Models and to test the performance of a regime-based asset allocation strategy which aims to capture returns based on the characteristics of flow-defined regimes. In an out-of-sample test, the strategy shows that it is possible to successfully capture returns by applying a regime-switching model to international portfolio flow data.

Acknowledgment

I would like to thank State Street Global Markets (SSGM) for providing the international portfolio flow data used in this project.

Introduction

This paper explores the effectiveness of a regime switching model constructed using international portfolio flow data, specifically cross-border equity flows, for capturing excess returns through dynamic asset allocation based on regimes.

Regimes are periods of time with unique characteristic financial variables, such as mean returns, correlations and volatilities. A change in regimes implies a change in the characteristic behaviour of the financial market which may continue for some time if the regime is persistent. In this case, hidden Markov Models (HMMs), an unsupervised machine learning technique is used to infer regimes. Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labelled responses.

Cross-border equity flows are composed of flows by shareholders, who add or remove cash from funds, and flows by managers, who buy or sell individual securities with fund deposits. Cross-border equity flows have been shown to be stable and persistent in nature. They have also been shown to influence equity returns.

In this paper regimes are considered as periods with a characteristic cross-border equity inflows or outflows. Through these regimes, the relationship between flows and equity returns is examined. If it is possible to infer stable, persistent regimes based on flows, it may be possible to capture returns based on the relationship between flows and returns.

A number of contributions to the literature are made in this paper. Firstly this paper shows that a HMM trained using cross-border equity flow data infers persistent, stable regimes. This paper also shows that flow-inferred regimes demonstrate characteristic mean returns. Combining these two findings, a portfolio is constructed using a regime switching based dynamic asset-allocation strategy. This strategy is shown to outperform a benchmark static equal weight buy and hold portfolio.

Literature Review

Regime switching models have gained popularity in quantitative finance since they were first introduced in Hamilton's (1989) seminal work which used a HMM to identify expansions and recessions in the business cycle. Regimes are an easy concept to grasp intuitively, a distinction can be made between "Bull" and "Bear" markets, periods of high and low volatility returns, as well as periods of change in policy or regulation (Ang & Timmermann, 2012). Regime switching HMMs have been shown to capture persistent regimes. The ability to infer regime changes has been shown to produce profitable dynamic asset allocation strategies which utilize regime switching. Ang and Bekaert (2002) concluded that a high-volatility, high-correlation regime is present in a regime-switching model is present during a bear market. Following on from this finding, they demonstrated that a regime switching based dynamic asset allocation strategy outperforms a static strategy by switching to cash during in high-volatility regimes (Ang & Bekaert, 2004). Bulla, Mergner, Bulla, Sesboüé, and Chesneau (2011) demonstrated that a regime-based asset allocation strategy under realistic assumptions could outperform a buy and hold strategy after taking transaction costs into account. More recent papers have demonstrated similar positive results using dynamic asset allocation strategies, using a HMM with time-varying parameters (Nystrup, Hansen, Madsen, & Lindström, 2015), (Nystrup, Hansen, Larsen, Madsen, & Lindström, 2017).

International portfolio flows have been shown to affect equity prices in developed and emerging markets. Cross-border equity inflows and outflows tend to cause international prices to rise and fall respectively (Tesar & Werner, 1994, 1995), (Brennan & Cao, 1997). Froot, O'Connell, and Seasholes (2001) found that flows to appear stationary and more persistent than returns, while also finding that flows have an influence on returns. Froot and Ramadorai (2008) show that weekly cross-border equity flows forecast emerging market equity returns.

Hidden Markov Model

The main characteristic of a hidden Markov model is a probability distribution of the observation $X_t,\ t = 1, \dots, T$ which is dependent on the states S_t of an unobserved first-order Markov chain.

A sequence of discrete random variables $\{S_t: t \in \mathbb{N}\}$ is said to be a first-order Markov chain if, for all of $t \in \mathbb{N}$, it satisfies the Markov property

$$\mathrm{P}(S_{t+1}|S_t, \dots, S_1) = \mathrm{P}(S_{t+1}|S_t) \tag{15}$$

A transition probability matrix (TPM) governs the switching behaviour of the model between states. In a two-state model for example, the TPM would be of the form

$$\Pi = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \quad (16)$$

with $p_{ij}, i, j \in \{1,2\}$ denoting the probability of being in state j at time $t+1$, given a sojourn in state i at time t. The observation X_t has a distribution at time t specified by $\mathrm{P}(X_t = x_t | S_t = s_t)$, the conditional or component distributions of the model. A two-state Gaussian component distribution would give:

$$x_t = \mu_{s_t} + \epsilon_{s_t}, \epsilon_{s_t} \sim N\left(0, \sigma_{s_t}^2\right) \quad (17)$$

Where $\mu_{s_t} \in \{\mu_1, \mu_2\}$ and $\sigma_{s_t}^2 \in \{\sigma_1, \sigma_2\}$. To estimate the parameters of the HMM an expectation-maximization (EM) method such as the Baum-Welch algorithm is commonly used (Baum, Petrie, Soules, & Weiss, 1970).

Assuming that successive observations are independent, the likelihood function is given by:

$$L(\theta) = \boldsymbol{\pi} \boldsymbol{P}(x_1) \Pi \boldsymbol{P}(x_2) \Pi \dots \boldsymbol{P}(x_{T-1}) \Pi \boldsymbol{P}(x_T) \mathbf{1} \quad (18)$$

where $\boldsymbol{P}(x_t)$ is a diagonal matrix containing the state-dependent conditional distributions as entries and $\boldsymbol{\pi}$ denoting the initial distribution of the Markov chain. After parameter estimation of the HMM, the hidden states can be inferred. A common technique for determining the most likely sequence of states is the Viterbi algorithm (Viterbi, 1967). The algorithm calculates the most probable sequence of states using:

$$\{\hat{s}_1, \dots, \hat{s}_T\} = \underset{j_1, \dots, j_T}{\operatorname{argmin}} P(S_1 = j_1, \dots, S_T = j_T | X_1^T = x_1^T) \quad (19)$$

Data

International equity flow indicators

The flow data used in this paper is part of a series of Active/Benchmark Equity Flow Indicators provided by SSGM, which are composed of flows by shareholder, who add or remove cash from funds, and flows by managers, who buy or sell individual securities with fund deposits.

The indicators are proprietary measures of investor behaviour developed by SSGM which represent the flows of aggregated portfolios.

The eight countries and regions used in the model are shown in Table 1. Cross-border equity flow indicator data corresponding to the eight countries and regions was obtained from SSGM for the time period January 2012 to December 2018 on a daily frequency.

Table 1 – Investment Universe

Investment Universe
Europe ex UK
EM Latin America
EM EMEA
Pacific ex Japan
EM Asia
USA
UK
Japan

Price data

MSCI Total Return index data was obtained from Thomson Reuters Datastream for the corresponding country and regional indices shown in Table 1 for the time period January 2012 to December 2018.

The daily 1-month USD LIBOR was obtained from Thomson Reuters Datastream for the time period January 2012 to December 2018. This was used as the risk-free rate when calculating the Sharpe ratio for the portfolio.

Methodology

A HMM was used to infer regimes based on cross-border equity flow data for the eight regions in Table 1. To characterise the regimes based on price movement, mean returns were calculated for each regime using MSCI Total Return Index data corresponding to the eight regions. To examine the ability of flow-based regimes to capture returns, a portfolio was constructed with the eight MSCI indices as assets. A long or short position was taken in each asset based on the mean return for current regime according to the HMM. If flows, and through them, regimes are persistent, the positions will capture consistent returns.

To initially estimate the parameters of the HMM and to characterise the inferred regimes based on their mean returns, a training period of January 2012 to December 2016 was chosen. The Viterbi algorithm was used to infer the most likely sequence of regimes up to the current regime.

Beginning with the regimes characterised in the 2012-2016 training period, an out-of-sample test was implemented from January 2017 to December 2018 using the following method which stepped through the data iteratively to exclude future information.

Using flow data from January 2012 up until time t, calculate the most likely sequence of regimes up until the current regime k.

The mean return is calculated for every instrument in the portfolio across the time periods spent in the current regime:

$$\mu_{t,k}^{i} = \frac{\sum_{t,k=1}^{T_k} r_{t,k}^{i}}{T_k} \quad (20)$$

Where $\mu_{t,k}^{i}$ is the mean return of instrument i at time t in regime k, T_k is the time spent in regime k and $r_{t,k}^{i}$ is the return of instrument i at time t in regime k.

Based on whether the mean return $\mu_{t,k}^{i}$ is positive or negative for the current regime k at time t, a simple long/short trading signal M_t^i is assigned to each instrument:

$$M_t^i = sign(\mu_{t,k}^{i}) \quad (21)$$

Each instrument in the portfolio is assigned an equal weight. To try and simulate a more realistic trading scenario the signal for time t is used to determine each weight at time $t+\tau$, where τ indicates a time shift. The weights are calculated as follows:

$$w_{t+\tau}^{i} = \frac{1}{N} M_t^i \quad (22)$$

Where $w_{t+\tau}^{i}$ is the weight of instrument i held in the portfolio at time $t+\tau$, N is the number of instruments in the asset class and M_t^i is the regime-based signal from the previous time period.

The return of the portfolio $r_{t+\tau}$ is calculated as:

$$r_{t+\tau} = \sum_{i=1}^{N} (r_{t+\tau}^{i} \, w_{t+\tau}^{i}) \quad (23)$$

Where $r_{t+\tau}^{i}$ is the return of instrument i at time $t+\tau$.

Results for the performance of a 4-regime model are included in Table 6. This number of regimes was chosen for stability reasons discussed in Appendix 2

Results

Estimating the parameters of a HMM using an EM algorithm gives the transition probability matrix (Equation 2). A TPM with a strong diagonal ($\Pi_{\mathrm{ii}} \gg 0.5$) indicates that the regimes will be highly persistent i.e. that the regime predicted at t will most likely be the regime predicted for $t+1$.

Training a 4 regime HMM over the period January 2012 to December 2018 using cross-border equity flow data, gives a TPM with a very strong diagonal ($\Pi_{\mathrm{f}_{ii}} \gg 0.9$).

$$\Pi_{\mathrm{f}} = \begin{pmatrix} 0.952 & 0.000 & 0.027 & 0.021 \\ 0.002 & 0.969 & 0.018 & 0.012 \\ 0.017 & 0.041 & 0.919 & 0.024 \\ 0.013 & 0.023 & 0.018 & 0.946 \end{pmatrix}$$

When a HMM is trained over the same period January 2012 to December 2018 using log returns calculated from the MSCI Total Return index data for the eight regions, gives a TPM with a much weaker diagonal. While $\Pi_{\mathrm{r}_{ii}} \gg 0.5$ for three of the regimes, indicating that they are persistent, the TPM of the HMM trained on flows is much stronger.

$$\Pi_{\mathrm{r_{ii}}} = \begin{pmatrix} 0.440 & 0.075 & 0.409 & 0.077 \\ 0.088 & 0.665 & 0.232 & 0.015 \\ 0.182 & 0.096 & 0.0712 & 0.010 \\ 0.124 & 0.022 & 0.012 & 0.842 \end{pmatrix}$$

The descriptive statistics in Appendix 1 for both daily returns and flow across the instruments in show clear distinctions between regimes over the period January 2012 to December 2018.

Out-of-sample testing was performed in an iterative fashion to obtain results which were as realistic as possible. The two versions of the model were trained with daily data up until time t used to determine the portfolio positions at time $t+\tau$. The first model has $\tau = 1$, with the predicted regime at time t determining portfolio positions 1 day ahead. The second model has $\tau = 3$, with the predicted state at time t determining the portfolio positions at 3 days ahead. The performance of the regime-based strategy was compared to a static equal weight portfolio containing the same instruments in the regime-based portfolio, as well as the MSCI World Index.

Table 6 shows key performance statistics - annualized return (AR), annualized volatility (Vol) and Sharpe Ratio (SR) - generated by the trading strategy in an out-of-sample test from January 2017 to December 2018.

The performance statistics are compared to a buy and hold strategy and to the MSCI World Total Return Index for the same period.

Table 6 – Performances of strategies and indices from Jan 2017 to Dec 2018, daily data.

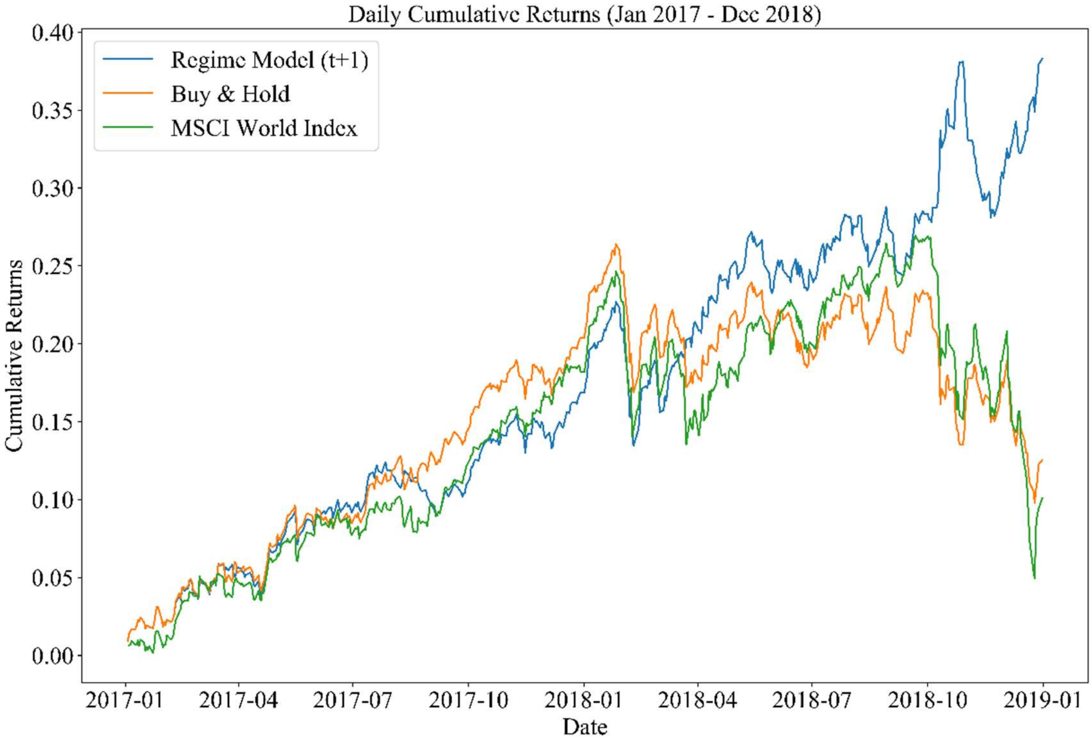
Strategy	**AR (%)**	**Vol (%)**	**SR**
Regime Model ($\tau = 1$)	16.23	7.55	1.84
Regime Model ($\tau = 3$)	14.97	7.46	1.69
Buy & Hold	5.89	8.07	0.45
MSCI World Index	4.77	9.75	0.25

The 1-day lag Regime Model realized the highest AR and SR. The 3-day lag Regime model had a slightly lower AR and SR but both strategies outperformed the benchmarks used for comparison across all performance indicators. The strong performance of the 3-day lag model demonstrates that returns captured using flow-based regimes are persistent, even if there is a delay in adjusting portfolio positions after the regime change.

The daily cumulative returns graph in Figure 1 shows the ability of the regime model to identify periods where there is a switch to a regime with negative mean returns and use this information to take a short position.

Appendix C shows the switching behaviour of the regimes in the out-of-sample test.

Figure 1 - Daily Cumulative Returns (January 2017 – December 2018)



Conclusions

Based on the results presented in this paper, a number of conclusions can be drawn.

The first finding of this paper is that the results indicate that regimes inferred using flow data are persistent, and that they are more persistent and stable than regimes inferred using returns, for the period January 2012 to December 2018 in the instruments examined ($\Pi_{\mathrm{f}_{ii}} \gg \Pi_{\mathrm{r}_{ii}}$)

It was also found that each regime inferred using flow data exhibited mean returns (Appendix A). Using this information along with the persistent nature of flow-based regimes, it was possible to persistent returns by adjusting portfolio positions to reflect switching regimes, taking advantage of the characteristic mean returns associated with each regime.

References

Ang, A., & Bekaert, G. (2002). International Asset Allocation With Regime Shifts. The Review of Financial Studies, 15(4), 1137-1187. doi:10.1093/rfs/15.4.1137

Ang, A., & Bekaert, G. (2004). How Regimes Affect Asset Allocation. Financial Analysts Journal, 60(2), 86-99. doi:10.2469/faj.v60.n2.2612

Ang, A., & Timmermann, A. (2012). Regime Changes and Financial Markets. Annual Review of Financial Economics, 4(1), 313-337. doi:10.1146/annurev-financial-110311-101808

Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. The annals of mathematical statistics, 41(1), 164-171.

Brennan, M. J., & Cao, H. H. (1997). International portfolio investment flows. The journal of Finance, 52(5), 1851-1880.

Bulla, J., Mergner, S., Bulla, I., Sesboüé, A., & Chesneau, C. (2011). Markov-switching asset allocation: Do profitable strategies exist? Journal of Asset Management, 12(5), 310-321. doi:10.1057/jam.2010.27

Froot, K. A., O'Connell, P. G. J., & Seasholes, M. S. (2001). The portfolio flows of international investors. Journal of financial economics, 59(2), 151-193. doi:https://doi.org/10.1016/S0304-405X(00)00084-2

Froot, K. A., & Ramadorai, T. (2008). Institutional Portfolio Flows and International Investments. The Review of Financial Studies, 21(2), 937-971. doi:10.1093/rfs/hhm091

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica: Journal of the econometric society, 357-384.

Nystrup, P., Hansen, B. W., Larsen, H. O., Madsen, H., & Lindström, E. (2017). Dynamic Allocation or Diversification: A Regime-Based Approach to Multiple Assets. The Journal of Portfolio Management, 44(2), 62-73. doi:10.3905/jpm.2018.44.2.062

Nystrup, P., Hansen, B. W., Madsen, H., & Lindström, E. (2015). Regime-Based Versus Static Asset Allocation: Letting the Data Speak. The Journal of Portfolio Management, 42(1), 103-109. doi:10.3905/jpm.2015.42.1.103

Tesar, L. L., & Werner, I. M. (1994). International equity transactions and US portfolio choice. In The internationalization of equity markets (pp. 185-227): University of Chicago Press.

Tesar, L. L., & Werner, I. M. (1995). Home bias and high turnover. Journal of international Money and Finance, 14(4), 467-492.

Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE transactions on Information Theory, 13(2), 260-269.

Appendix A – Descriptive Statistics (January 2012 – December 2018)

Table 3 – Mean daily return for each instrument during the periods spent in each regime

	Mean Daily Return (%)			
	Regime 1	**Regime 2**	**Regime 3**	**Regime 4**
Europe ex UK	0.027	0.031	-0.239	-0.015
EM Latin America	0.032	0.040	0.027	0.115
EM EMEA	0.016	0.042	-0.093	0.051
Pacific ex Japan	0.031	0.049	-0.140	0.034
EM Asia	0.038	0.015	-0.164	0.097
USA	0.095	-0.035	-0.470	0.072
UK	0.008	0.070	-0.252	-0.010
Japan	0.060	0.003	-0.370	-0.022

Table 4 - Mean annualized volatility for each instrument during the periods spent in each regime

	Annualized Volatility (%)			
	Regime 1	**Regime 2**	**Regime 3**	**Regime 4**
Europe ex UK	9.57	10.58	16.88	12.12
EM Latin America	13.26	12.80	13.83	16.32
EM EMEA	11.44	12.81	10.39	12.53
Pacific ex Japan	7.98	10.49	11.41	9.76
EM Asia	10.42	13.96	12.17	12.36
USA	10.15	13.88	17.06	14.77
UK	9.36	11.01	14.62	11.47
Japan	13.26	15.07	17.45	14.43

Table5 – Mean daily cross-border equity flow for regime during the test period

	Mean Daily Cross-Border Equity Flow			
Country/Region	**Regime 1**	**Regime 2**	**Regime 3**	**Regime 4**
Europe ex UK	0.037	-0.023	0.021	0.018
EM Latin America	0.107	0.037	-0.160	0.018
EM EMEA	0.066	0.009	0.002	0.026
Pacific ex Japan	0.039	-0.035	-0.019	0.011
EM Asia	0.032	0.017	0.012	0.012
USA	-0.057	-0.133	-0.040	-0.008
UK	0.080	-0.086	0.000	0.021
Japan	0.027	-0.045	-0.025	0.009

Table 6 - Daily cross-border equity flow standard deviation for each regime during the test period

	Daily Cross-Border Equity Flow Standard Deviation			
Country/Region	**Regime 1**	**Regime 2**	**Regime 3**	**Regime 4**
Europe ex UK	0.082	0.091	0.062	0.029
EM Latin America	0.141	0.144	0.146	0.070
EM EMEA	0.077	0.071	0.079	0.046
Pacific ex Japan	0.049	0.066	0.042	0.029
EM Asia	0.053	0.084	0.075	0.039
USA	0.082	0.081	0.049	0.043
UK	0.087	0.166	0.069	0.043
Japan	0.058	0.061	0.079	0.031

Appendix B – Model Stability

As the HMM is trained and its parameters estimated in an unsupervised fashion when identifying regimes, there can be some variation in the regimes assigned to each day. To test the performance and stability of the model over a number of regimes, the model testing and training was repeated 50 times for each model to determine the stability of the results. The 95% confidence interval for the annualized return figure shows that the expected result varies as the number of regimes increases.

Table 7 - Daily Model Performance (January 2017 – December 2018)

	4 Regime Model	**5 Regime Model**	**6 Regime Model**	**7 Regime Model**	**Buy & Hold**	**MSCI World Index**
Annualized Returns (%)	**16.23**	**15.53**	**16.88**	**16.22**	5.89	4.77
95% Confidence Interval (±%)	1.88	1.89	2.41	3.04	-	-
Volatility (%)	7.55	7.12	7.31	7.15	8.07	9.75
Sharpe Ratio	1.84	1.84	2.00	1.89	-	-

Appendix C – Out-of-Sample Test Regime Switching

Figure 2 – The percentage of time spent in each regime during the out-of-sample test. Regime 1 was the most persistent regime.

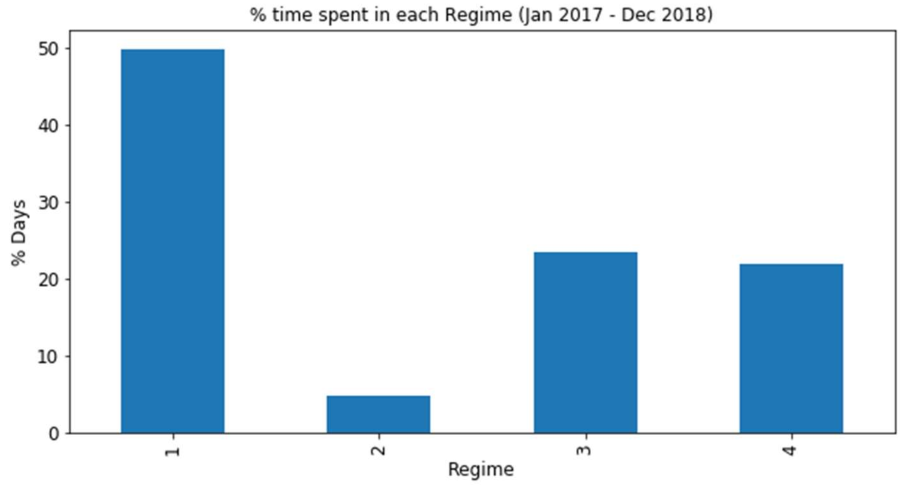


Figure 3 – Switching behaviour between regimes during the out-of-sample test.