TILBURG UNIVERSITY

MASTER THESIS

The advantages of hidden Markov models for Dutch pension funds.

BY: WOUT TEMMINK

SUPERVISOR: BERTRAND MELENBERG

Tilburg University

TiSEM - Tilburg School of Economics and Management

March 13, 2024

Contents

Abstract

Dutch pension funds could improve their investment policies by integrating more tactical asset allocation (TAA) components into their existing strategic asset allocation (SAA) framework. Hidden Markov models (HMMs) are demonstrated to be more effective in capturing stylized effects in the distribution of asset returns such as volatility clustering, fat-tails, and skewness than the linear models currently used by major Dutch pension funds. Moreover, empirical analysis demonstrates that HMMs have the capability to generate regime forecasts that are statistically significant in comparison to forecasts obtained from random guessing. These forecasts have the potential to generate excess risk-adjusted returns above the benchmarks. Evident from extensive empirical analysis of HMM-based investment strategies using real-world data.

The statistical significance of the HMM's ability to forecast asset returns is tested and shown to be insignificant. Furthermore, there is no concrete proof that these return forecasts can deliver excess economic value when compared to benchmarks in terms of risk-adjusted returns. Through the integration of statistical and economic aspects in the evaluation of regime and return forecasts, this research also gives insights into the positive relationship between statistical accuracy and economic value. Lastly, consideration is given to the practical implications of implementing HMM-based investment strategies at Dutch pension funds. It is concluded that major changes to the governance structure of the funds are needed for the successful implementation of these strategies.

Chapter 1

Introduction

1.1 Introduction

In light of the new pension system mandated by the Wet Toekomst Pensioenen (WTP) law in July 2023, Dutch pension funds should reassess their investment policies. Considering that under the new pension system, investment performance has a more direct impact on pension outcomes. Altering the investment policy of an average Dutch pension fund is a complex and extensive process. Consequently, this study will focus on the strategic asset allocation (SAA) of the fund. The long-term horizon of its liabilities has a significant impact on the traditional SAA approach that the typical Dutch pension fund uses in the country's current pension system. This study aims to reevaluate the SAA of Dutch pension funds in order to achieve an increase in risk-adjusted returns for the participants of the fund, while adhering to the constraints of the new pension system.

Prior to delving further into alternative asset allocation models, it is important to provide more background information on the issue at hand. This is accomplished by an examination of the former and current pension systems, leading to the formulation of the problem statement and a set of research questions that will guide the process of finding a solution to the problem statement. First, the problem statement: Dutch pension funds need to reevaluate their way of determining their investment policy by incorporating more tactical asset allocation (TAA) elements into their SAA in order to provide the participants of the fund with the best possible service.

The first step undertaken to address this issue is a literature review. This literature review aims to get a comprehensive overview of the historical development of asset allocation and also to provide a cross-section of some of the effective asset allocation strategies documented in the literature. It is evident from this discussion that the modern portfolio theory (MPT) proposed by Markowitz (1952) lays the foundation for many of the asset allocation models that can be found in the literature. Moreover, it continues to be utilized by many of the major Dutch pension funds. Nevertheless, from this literature review, it is also evident that MPT has its limitations, primarily due to the underlying assumption that an investor's asset allocation is solely based on the risk and return of the assets being studied. While this assumption presents significant issues, like the fact that stylized effects such as volatility clustering, fat tails, and skewness in the distribution of asset returns cannot be incorporated, it is also very interpretable and leads to straightforward mathematical calculations. Due to these considerations, the decision is made to explore alternative ways of modeling asset return distributions to address the restrictions while maintaining the advantages of MPT. More specifically, a model is chosen that is best able to capture stylized effects present in the asset return distribution while also being interpretable. The most appropriate model proved to be the hidden Markov models (HMMs). One advantage of using HMMs when modeling asset returns is that the model assumes that there are multiple regimes underlying the asset return distribution, which aligns with economic reasoning. Another advantage is that this way of modeling has advantageous characteristics, as it allows us to capture stylized effects in the distribution of asset returns (Ang and Bekaert, 2002; Guidolin and Timmermann, 2008). After discussing the HMM's potential to improve the asset allocation of Dutch pension funds through the remodeling of asset return distributions, the next step is to obtain data with which this hypothesis can be verified. I define an investment universe that is similar to that of a Dutch pension fund. Furthermore, a set of macroeconomic variables is defined that will be used to determine the various regimes in the stock market.

Prior to proceeding with testing the models and obtaining results, it is important to provide a comprehensive overview of the HMM itself. Chapter 4 provides an overview of the theory needed for the implementation of HMMs. This includes details on regime probability estimation, forecasting, initialization, and parameter estimation. Additionally, due to the modular nature of the HMM, this section also includes information on which models are interesting for research.

Chapter 5 presents the empirical results of the model fit analysis. The results clearly demonstrate that HMMs outperform their linear counterparts. Therefore, it is advisable to model asset return distributions according to the aforementioned class of HMMs. Furthermore, it provides information on how well the HMM captures the stylized effects in the asset return distribution. Once the evaluation of the model fit is completed, the next step is evaluating the regime forecasting potential of the HMM. The HMM's regime forecasting performance is evaluated based on its statistical accuracy and the economic value it generates. The analysis shows that the model is able to produce statistically significant results while also providing excess economic value when compared to benchmark models that do not utilize HMMs. These regime forecasts enable me to also forecast asset returns. Moreover, these forecasts are also evaluated based on their statistical accuracy and economic value. From the analysis, it is evident that the HMM does not produce statistically significant results or provide economic value.

All of the findings are outlined in Chapter 6. Additionally, recommendations for improving risk-adjusted returns and optimizing asset allocation strategies within the new pension system are given. Furthermore, Chapter 7 discusses some of the practical difficulties that may arise during the implementation of such asset allocation strategies. The practical implications of adopting HMM-based investing strategies are enormous, as they require substantial modifications in the fund's governance structure. In addition, potential areas for further research are discussed, such as the possible incorporation of environmental, social, and governance (ESG) factors.

1.2 Motivation problem

Dutch pension funds are among the biggest investors in the world, with their total capital under management growing to a staggering €1.736 billion during the year 2023 (DNB, 2023). With around 20% of their total capital invested in listed stocks, they are also a prominent participant in the (inter)national stock market. Dutch pension funds outsource a lot of their day-to-day activities to external consulting firms because of efficiency considerations. One of such consulting firms is Sprenkels. Sprenkels was originally founded in 1999, and they have been one of the leading consulting firms in the Dutch pension market for over 15 years. Sprenkels' primary goal is to offer its clients the best possible advice, and a crucial aspect of achieving this is staying up-to-date on industry trends and developments.

Sprenkels' consulting services include, but are not limited to, actuarial, investment, risk management, communication, and legal advice. One of its core activities is providing clients, mostly Dutch pension funds, with asset liability management (ALM) studies. An ALM study is an exhaustive analysis of the pension fund's assets and liabilities. The goal of the study is to ensure that the fund can meet its long-term obligations to pensioners. Typically, Sprenkels provides its clients with an ALM study every three years. One key element of an ALM study is evaluating the client's investment policy. Evaluating an investment policy involves, among others, determining the weights of different asset classes within the investment world of the client, which is also called determining the strategic asset allocation (SAA). There are many factors that need to be considered when it comes to the process of determining weights, like regulations, environmental, social, and governance (ESG) goals, future liabilities, and desired returns. The SAA of a pension fund will be based to a large degree on its long-run liabilities. This means that for the development of asset allocation strategies, a long horizon of at least 15 years is taken into account. Another important factor influencing the development of asset allocation strategies is the dynamics of the current pension system. Most of the large pension funds are structured following a collective pension system, where, through a mechanism called the "doorsneesystematiek," risks and benefits are shared throughout the generations.

Due to the aforementioned long horizon and the nature of the current pension system, not much attention is being paid to the potential influence of the business cycle on asset return distributions. Consequently, less attention is being paid to the potential short-term effects of the business cycle on asset allocation and fund performance. The fact that less attention is being paid to short-term effects runs counter to the fact that, in addition to the three year ALM study, a client can also choose to increase the frequency of such studies, requiring Sprenkels to perform ALM studies annually. These studies primarily focus on revisions to the current investment policy. In particular, the goal of these annual studies is to check whether the fund remains on track while also making reactionary adjustments, including revising clients' investment policies. However, due to the static nature of the models currently used and the aforementioned long horizon, these annual studies rarely result in major adjustments, which in certain cases might be warranted.

This problem becomes more problematic in light of the enactment of the new pension system mandated by the Wet Toekomst Pensioenen (WTP) law on July 1, 2023. With the introduction of the WTP, there is now a renewed interest in the potential significance of the business cycle for the asset allocation strategies and returns of Dutch pension funds. This renewed interest is a consequence of the fact that two key mechanisms will be less commonly observed in the switch to the new system. Firstly, with the abolishment of the "doorsneesystematiek," every participant now accumulates their own personal pension wealth based on the premium they pay every month and their age-dependent life cycle. Secondly, pension benefits are no longer guaranteed but are linked to the performance of the underlying investment portfolio. This means that now, more than in the current system, not only the SAA and the long-term performance of a portfolio matter but also the tactical asset allocation (TAA) and the medium-term portfolio performance.

Prior to formally identifying the problem, the distinctions between SAA and TAA will be highlighted. As noted by Eychenne et al. (2011) there is a clear difference between SAA and TAA, namely the time horizon considered during the decision process, as illustrated in Figure 1.1. SAA typically considers a time window of 10 to 50 years, while TAA is more short-term, focusing on a time window of three months to three years. Lastly, Eychenne et al. defined the concept of market timing (MT), which considers an even smaller time window. MT will not be relevant for the types of investors that are considered in this research since this type of trading does not fit the risk appetite of the average participant in a pension fund. Furthermore, the governmental regulations that are put in place would also not allow it.

Figure 1.1: Types of asset allocation, sorted by the length of the time window considered (Eychenne et al., 2011).

Differences in time windows considered lead to variations in asset allocations. An example of such a difference could be that the TAA of an investor is more heavily invested in domestic stocks than the SAA of that very same investor, at the cost of a smaller allocation to corporate bonds. Campbell and Viceira (2002) argue that these differences between the TAA and the SAA of an investor are a result of short-term fluctuations in risk premia. It is a well-known fact that risk premia fluctuate over the business cycle; their movements are directly influenced by an array of macroeconomic factors. Macroeconomic factors that make up the business cycle are, for example, the prevailing inflation rate, the unemployment rate, and GDP growth. Furthermore, policy tools can also affect the business cycle and thus risk premia. Hayo et al. (2010); Anzuini et al. (2012); Joyce et al. (2010) have all shown that monetary policies of central banks, which drive and simultaneously are driven by the business cycle, can significantly influence asset returns. Based on the information provided in this section, it can be inferred that there is merit to including TAA elements in the investment strategy of an institutional investor. Moreover, the implementation of the new pension system will amplify these advantages.

1.3 Problem definition

From the conclusion of Section 1.2, it is clear that Dutch pension funds need to reevaluate their way of determining their investment policy by incorporating more TAA elements into their SAA, especially in light of the introduction of the new pension system. Properly evaluating the entire process of determining an investment policy involves too many factors to discuss in a single study. Hence, to maintain a manageable scope for this study, the primary focus of this study will be to optimize the SAA of an institutional investor, measured in terms of risk-adjusted returns generated by their investment portfolio, by incorporating elements of TAA. Performance factors like environment, social, and governance (ESG) scores are therefore purposefully not taken into account.

Research Questions

1. Which model is the most efficient in capturing stylized effects commonly observed in the return distributions of an institutional investor portfolio over time?

2. Are HMMs able to produce stock market regime forecasts and asset return forecasts that are significantly more accurate than random guessing?

3. Can HMM-based asset allocation models significantly outperform their standard linear-based counterparts in terms of risk-adjusted return in the context of an institutional investor like a Dutch pension fund?

To address these three research questions, a comprehensive literature review will be conducted in Chapter 2. Subsequently, all the data that will be utilized in this study will be outlined in Chapter 3. Chapter 4 will present the outline of HMMs and the approach used to address these questions. Then, in Chapter 5, a detailed discussion of the results will be provided. Finally, Chapters 6 and 7 will address the research questions and offer guidance for future research. By implementing these actions, I offer Sprenkels a thorough comprehension of some the possible benefits that can be obtained in the new pension system.

Chapter 2

Literature review

In Section 2.1, I provide an overview of the history of asset allocation and a cross-section of possible asset allocation models. A majority of the institutional investors considered in this study use modern portfolio theory (MPT) to some extent. Therefore, it is decided to discuss the key assumptions and limitations of MPT in Section 2.2. The limitations of the MPT approach suggest that it may not be the optimal asset allocation strategy. These considerations raise the question: is it possible to uphold the core principles of the MPT approach and retain its useful mathematical properties while addressing its limitations of. The models presented in Section 2.3 aim to capture the stylized facts discussed in Section 2.2 more effectively than the benchmark model while still staying within the constraints of MPT. To conclude, in Section 2.4 I will select the most promising model.

2.1 Introduction to asset allocation

Much of the literature addressing asset allocation can be traced back to the seminal work of Markowitz (1952). Mean-variance optimization, as introduced by Markowitz is one of the simplest and most effective approaches used to obtain optimal weights for the various assets in an investor's portfolio (Otranto, 2010). Mean-variance optimization simplifies the asset allocation problem to a constrained maximization problem, where the investor chooses weights for all of the available asset classes such that the risk-adjusted expected return of the portfolio is maximized and the constraints on the portfolio choice are respected. This implies that investors' decisions are determined solely by the mean and the correlation of the asset return distributions.

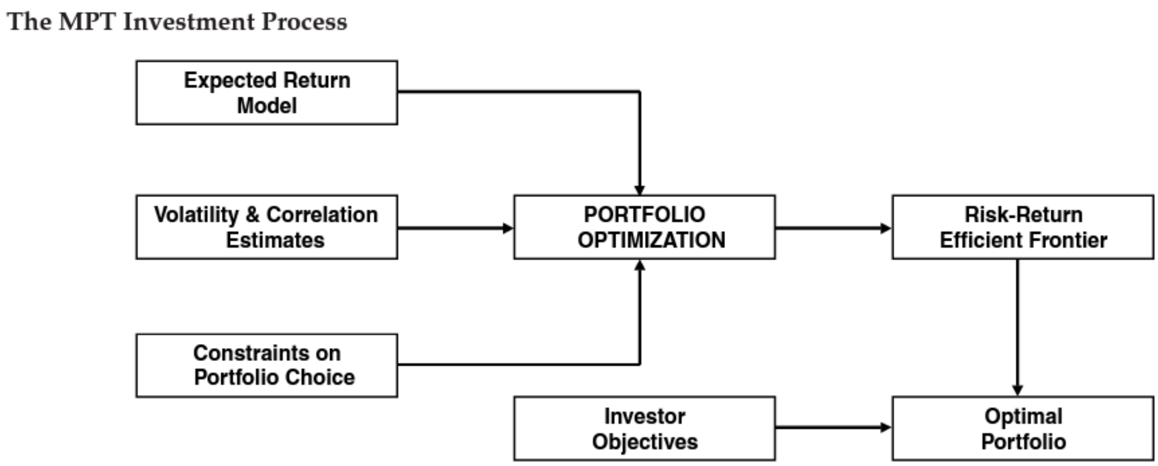

Figure 2.1: The MPT Investment process (Fabozzi et al., 2002)

The framework in which mean-variance optimization is frequently used is called modern portfolio theory (MPT). Figure 2.1 displays a clear overview of the investment process of an MPT investor. In their paper, Fabozzi et al. (2002) clearly describe the process an MPT investor goes through in order to find the optimal investment portfolio. Firstly, the investor determines a set of asset classes that will be considered in his investment universe. Secondly, estimates of mean returns and correlations are determined, typically by analyzing historical data. These estimates are then used as inputs for the portfolio mean-variance optimization. The output of this optimization is an efficient frontier, which is defined as the maximum expected return of a portfolio at each level of portfolio volatility from the selected asset classes (Hodnett et al., 2012). An example of such an efficient frontier can be found in Figure 2.2.

The efficient frontier represents many efficient portfolios; in order to find the optimal portfolio, one needs the capital asset pricing model (CAPM). The CAPM was introduced by Sharpe (1964), and it is still one of the most popular methods used for estimating expected returns by decomposing the expected return of an asset into the risk-free rate and a risk premium. The risk premium can be further decomposed into the correlation of the asset with respect to the market portfolio and the market risk premium. An important assumption of the CAPM is that all investors are rational and will therefore invest in a portfolio situated on the Capital Market Line (CML). Investing on the CML implies making the optimal trade-off between risk and return, in equilibrium, by borrowing or lending at the risk-free rate (Sharpe, 1964). The CML is a special case of the Capital Allocation Line (CAL). The CAL represents all the combinations of an investment portfolio and the risk-free asset. The slope of the line is called the Sharpe ratio. As stated, the optimal CAL is called the CML, and under the assumptions of CAPM, it offers the highest possible expected return for any given level of risk and the lowest possible risk for any given level of expected return (Hodnett et al., 2012). Examples of both the CML and various CALs can be found in Figure 2.2. The optimal portfolio is then defined as the intersection between the CML and the efficient frontier. Alternatively, the optimal portfolio can be found by using Monte Carlo simulations; this method computes wealth distributions for the possible portfolios on which the ultimate portfolio decision is based. Naturally, an important factor dictating the outcomes of MPT are the estimates of the mean returns and correlations.

A commonly used variation of mean-variance optimization is global minimum variance optimization, where instead of looking at mean returns and correlations, one only looks at correlations. In Figure 2.2 one can identify the resulting portfolio of global minimum variance optimization as the minimum variance portfolio situated on the efficient frontier.

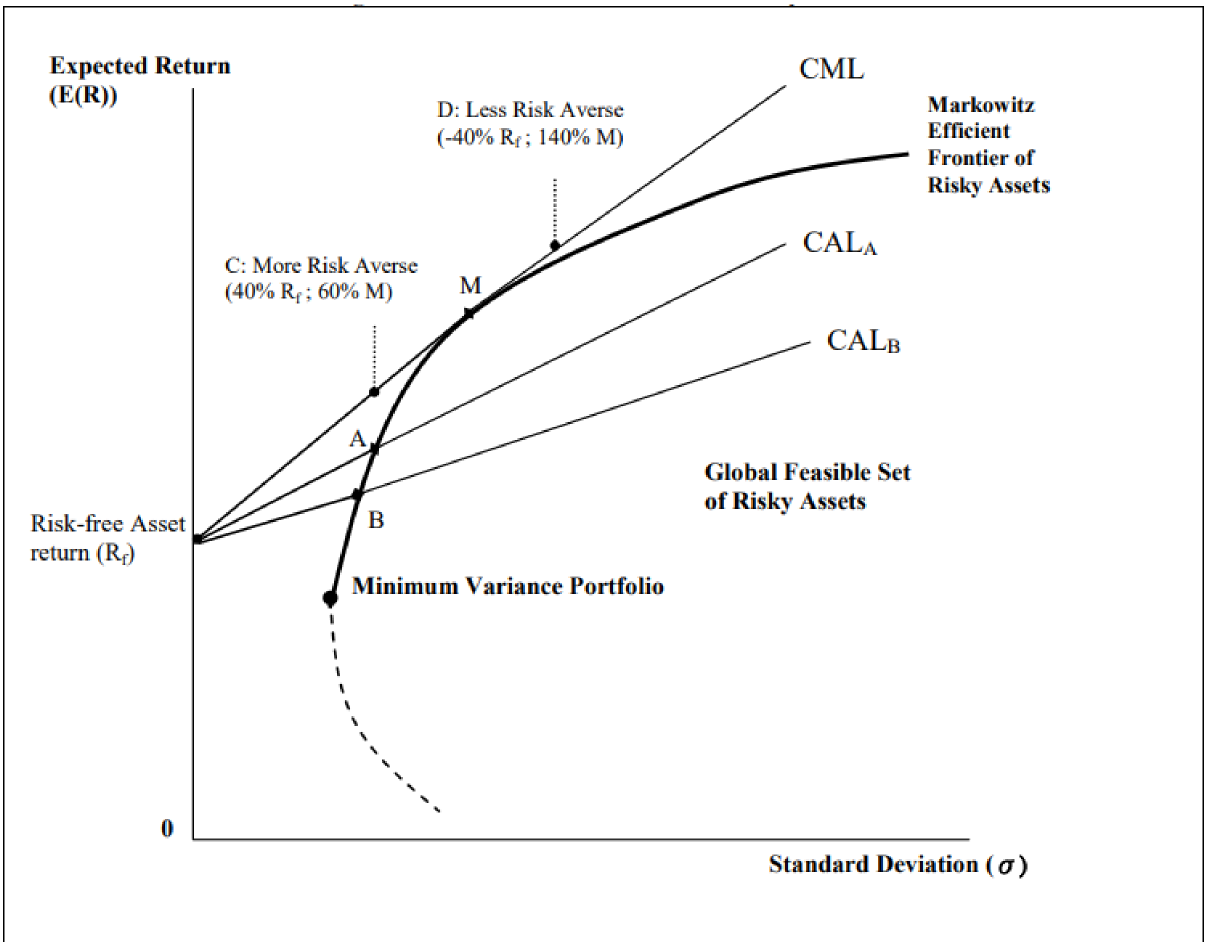

Figure 2.2: Markowitz efficient frontier (Hodnett et al., 2012)

Notes: This figure displays an example of a Markowitz efficient frontier. The graph includes multiple Capital Allocation Lines (CALs) as well as the Capital Market Line (CML). M is the optimal portfolio, it is located at the intersection of the CML and the efficient frontier. Furthermore, the minimum variance portfolio is located on the efficient frontier where the standard deviation (σ) is smallest.

An extension of the Markowitz (1952) approach would be to allow investors to incorporate subjective beliefs about the direction of the stock market in their modeling. Black and Litterman (1992) were the first to extend the Markowitz framework by incorporating investors' subjective beliefs into their modeling. The Black and Litterman model extended on the Markowitz framework by combining the equilibrium CAPM with subjective opinions on asset returns (Walters et al., 2014). The approach of Black and Litterman is just one of the many examples of models belonging to the behavioral finance class of asset allocation models. In more recent work, Almgren and Chriss (2007) introduced an asset allocation model based on directional signals instead of point estimates of asset returns. While still relying on the framework of Markowitz (1952), they offered a model that, like Black and Litterman (1992), allows investors to incorporate their subjective opinions into the model. Moreover, the use of directional signals instead of point estimates as input for the optimization ensures that the model is not vulnerable to changes in point estimates of asset returns and, as a result, makes the model more robust. This is in contrast to the frameworks put forward by Markowitz (1952) and Black and Litterman (1992).

MPT produces static asset allocation strategies, implying that the asset allocation is determined at a single point in time. Consequently, the allocation remains fixed for a predefined period of time, and then the optimal asset allocation is reevaluated. This approach might not be appropriate when dealing with dynamic stock market returns. As a result, much of the literature has been focused on dynamic asset allocation. One of the first papers that tried to deal with the aforementioned issue is Merton (1969). Merton considered the dynamic nature of financial markets in his formulation of an optimal investment strategy. This was done by modeling asset returns as stochastic processes, thus introducing a time factor. In his formulation, he still considered the Markowitz mean variance trade-off between risk and return.

The CAPM is often referred to as the one-factor model since it only considers the market return as a 'factor' explaining an asset's return (Haugh, 2016). In addition to this simple one-factor model, there are extensions containing multiple factors, like the famous three-factor model developed by Fama and French (1993). Fama and French consider not only the market factor but also the outperformance of small-market cap companies relative to large-market cap companies and the outperformance of high book-to-market value companies versus low book-to-market value companies as factors to explain asset returns. Both the CAPM and the Fama-French three-factor model belong to the general class of factor investing.

Post modern portfolio theory (PMPT) is a broader framework of which the MPT is a special (symmetrical) case (Rom and Ferguson, 1994). PMPT was developed as a response to some of the limitations of MPT. PMPT is predicated on the assumption that risk is not symmetrical, as opposed to MPT. Risk is deemed to be asymmetrical in the sense that "PMPT's downside risk' measure makes a clear distinction between downside and upside volatility. In PMPT, volatility below the investor's target return incurs risk; all returns above this target cause "uncertainty," which is nothing more than a riskless opportunity for unexpectedly high returns" (Rom and Ferguson, 1994, p. 351). This target rate of return could, for example, be the risk-free rate.

A more recent development in the world of asset allocation is the introduction of machine learning. Machine learning techniques try to find predictability patterns in asset return data, as these can lead to risk-adjusted excess returns. The work of Turner and Han (2009) is one example of how more advanced machine learning techniques can be used to obtain improvements in risk-adjusted returns. More specifically, Turner and Han evaluates the possibility of formulating the portfolio optimization problem as a reinforcement learning problem, where fitted Q-iteration is used to find the optimal policy given each economic state. Of course, there are way more potentially interesting machine learning techniques that have been used to optimize asset allocation. However, the specific nature of Sprenkels' clients' business operations prevents the use of such models due to governance structures and regulatory constraints.

So far, I have explored a cross-section of the available asset allocation models. Each of them has their own advantages and disadvantages. Currently, most of the major Dutch pension funds utilize the MPT framework together with historical averages to estimate mean returns and correlations.

2.2 Limitations of MPT

In order for MPT to be effective, a set of assumptions has to hold. One of the most important assumptions underlying the MPT approach is that investors are assumed to base their investment decisions exclusively on the trade-off between risk, which is measured by the correlation between the asset return distributions, and return, which is measured by the mean of the asset return distributions. This assumption is highly controversial since it indirectly requires the joint asset return distributions to be normal. The reason normality is required is that the correlation is only a complete measure of risk if the joint multivariate distribution of asset returns is normal. That is, covariance is only an exhaustive measure of co-movement if the joint distributions are normal themselves. This implicit normality assumption is an issue since normally distributed asset return data is rarely observed in the real-world (Guo, 2022). Furthermore, Jondeau and Rockinger (2005); Ang and Timmermann (2012) show in their work that asset return distributions often display stylized facts like fat-tails and skewness. These non-linear patterns cannot effectively be captured by MPT, thus violating the aforementioned MPT assumption. Consequently, these violations imply that MPT may not accurately capture the risk and return characteristics of asset returns.

Furthermore, Fabozzi et al. (2002) have shown that estimating distribution parameters with historical data might lead to widely varying results depending on the time period considered. This supports the notion that there are time-varying effects in the distribution of asset returns. The aforementioned variation is caused by effects like volatility clustering. Volatility clustering was first discovered by Mandelbrot and Mandelbrot (1997). It describes the effect that periods of significant returns are followed by periods of, either positive or negative, significant returns. While periods with small returns are followed by periods with small returns. Stylized effects like volatility clustering, fat tails, and skewness cannot effectively be captured by a normal distribution. This implies that MPT requires assumptions that are unrealistic and, therefore, cannot work in every scenario.

Another highly controversial assumption made in MPT is that of market efficiency. This theory is called the efficient market hypothesis (EMH), and it was popularized by Fama (1970). It states that stock prices reflect all available information and that it is impossible to consistently outperform the market on a risk-adjusted basis. Proponents of the EMH conclude that, because of the EMH, the best possible outcome an investor can obtain is the market return. Consequently, institutional investors should focus on investing in a low-cost, passive portfolio. I acknowledge that there are more (controversial) assumptions required by the MPT framework that could be discussed in this paper. However, due to time constraints, the decision is made to tackle the validity of these assumptions.

2.3 Modelling of asset return distributions

2.3.1 Classical model

The estimation model that is most commonly used in the literature is the one that was first outlined by Markowitz (1952) (Fabozzi et al., 2002). In this model, mean returns and correlations of asset return distributions are estimated based on historical averages; for example, many of Sprenkels' clients take into account a horizon of 15 years for this.

2.3.2 Vector autoregressive models

The general class of linear predictability models, known as vector autoregression VAR(p) models is an expansion of the classic model. Much of the literature on long-run dynamic asset allocation under predictable returns has focused on linear predictability models (Guidolin and Hyde, 2012). VAR(p) models predict returns by looking at their lagged values and a set of exogenous variables; the number of lags considered is indicated by the value of p. Possible exogenous variables are dividend yield, unemployment, and the term spread. Linearity implies that a movement in one or more of the exogenous variables today will lead to a linear movement in returns tomorrow. VAR(p) models have obvious advantages, like the fact that they are relatively simple. However, they also have major drawbacks. For example, implementing linear models to analyze asset returns may not be appropriate due to the frequent observation of nonlinear patterns in asset returns (Ang and Bekaert, 2002; Guidolin and Timmermann, 2007; Guidolin and Hyde, 2012). The general VAR(p) model is specified as:

$$\begin{bmatrix} \mathbf{y}_{t+1} \\ \mathbf{x}_{t+1} \end{bmatrix} = \boldsymbol{\mu} + \boldsymbol{\beta} \begin{bmatrix} \mathbf{y}_{t} \\ \mathbf{x}_{t} \end{bmatrix} + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \text{iid}\, N(\mathbf{0}, \boldsymbol{\Omega}). \tag{2.1}$$

In equation 2.1, $\mathbf{y}_t$ is defined as the vector of asset returns at time t, and $\mathbf{x}_t$ is defined as the vector of exogenous variables at time t. Furthermore, $\boldsymbol{\mu}$ is defined as the intercept, and ε_t is defined as the random shock. Lastly, $\boldsymbol{\beta}$ is defined as the regression coefficient. Please note that by not including any exogenous variables and setting $\boldsymbol{\beta}$ equal to 0, the classical model is once again obtained.

2.3.3 Regime approach, one (composite) variable

One can extend further on the VAR(p) model and try to deal with time-varying effects by incorporating the regime approach of Brocato and Steed (1998). In their work Brocato and Steed, divide the asset return series into historical periods according to turning points in the National Bureau of Economic Research (NBER) business cycle. Brocato and Steed show that a strategy using cyclical allocation on the basis of the business cycle significantly outperforms a buy-and-hold benchmark portfolio in-sample. They classify every month between 1972 and 1993 as either an expansionary or contractionary month based on whether the NBER business cycle is contracting or expanding. Based on the aforementioned classification, estimates for the asset return distributions in both subsets of the data are found; these estimated distributions are then used to determine the optimal asset allocation. Furthermore, Brocato and Steed find that there are significant differences in the correlation matrices between regimes, which is further evidence for the idea that there are time-varying effects present in the distribution of asset returns.

Jensen and Mercer (2003) extend the research of Brocato and Steed (1998) by considering turning points in the monetary cycle. Their research follows the approach Brocato and Steed closely. They look at the monetary cycle, identify turning points, and use these turning points to divide the asset return series into regimes. Jensen and Mercer assume that there are two identifiable regimes in the monetary cycle; a regime can either be classified as expansionary or contractionary. In a similar manner to Brocato and Steed all time periods are aggregated into one of the two regimes, and distribution parameters are estimated on the basis of these time periods. The estimated distribution parameters are then used as inputs for the mean-variance optimization. The difference between the work of Jensen and Mercer and the work of Brocato and Steed is that the work of Jensen and Mercer allows for out-of-sample testing since the monetary cycle indicator uses only ex ante information.

Lastly, Dzikevičius and Vetrov (2012) uses the Organization for Economic Cooperation and Development (OECD) Composite Leading Indicator (CLI) to determine regimes in asset return series. The indicator is leading, which means it tends to change before the economy changes. This property makes it a useful forecasting tool. It finds regimes based on a composite leading indicator function; four states are defined based on whether the indicator is above or below the long-term trend and whether it is decreasing or increasing. Again, for every state, return distributions are estimated. Based on these, a mean-variance strategy can be used to find portfolio weights. All of the strategies described in this section assume that the distribution of returns is Gaussian. However, in contrast to the classical model, these approaches are able to deal with time-varying effects.

2.3.4 Hidden Markov models

The use of hidden Markov models (HMMs) is another effective method to improve asset return distribution estimation by including cyclical patterns of financial markets. Existing work strongly supports the idea that the business cycle greatly influences the distribution of asset returns and the ability of HMMs to capture these effects (Ang and Bekaert, 2002; Guidolin and Timmermann, 2008). HMMs include the aforementioned impact in the return distribution by introducing regime shifts (Guidolin and Timmermann, 2006, 2008). Each regime corresponds to specific stages in the business or financial cycle. Guidolin and Timmermann demonstrate that a model with four distinct regimes can effectively capture the dynamics of the joint distribution of returns. These regimes include a moderately persistent bear state, a highly persistent low-volatility bull state, a general bull state, and a highly volatile transient state. This way of modeling will introduce a new point of view from which return distributions and hedge requirements can be evaluated over time. The general form of these models can be found in equation (2.2).

$$\begin{bmatrix} \mathbf{y}_{t+1} \\ \mathbf{x}_{t+1} \end{bmatrix} = \boldsymbol{\mu}_{s_{t+1}} + \boldsymbol{\beta}_{s_{t+1}} \begin{bmatrix} \mathbf{y}_{t} \\ \mathbf{x}_{t} \end{bmatrix} + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \text{iid}\left(\mathbf{0}, \boldsymbol{\Omega}_{s_{t+1}}\right). \tag{2.2}$$

The notation in equation (2.2) is similar to the notation in equation (2.1), only now the variables $\boldsymbol{\mu}$,$\boldsymbol{\beta}$ and $\boldsymbol{\Omega}$ are dependent on a latent variable. The latent variable s_t is generally governed by a transition probability matrix that is assumed to be constant over time. Haase and Neuenkirch (2023) have shown in their work that incorporating macroeconomic predictor variables can lead to a more effective model in terms of statistical performance. The performance of these models has been tested by Guidolin and Hyde (2012), they evaluate the SAA problem of an individual investor with a utility function and constant relative risk aversion (CRRA). The investment portfolio is rebalanced every month, and an investor has a long-term horizon of 5 years. Guidolin and Hyde (2012) show in their research that the VAR(p) yields a lower realized performance than that of the considered HMMs. Furthermore, the HMM differs from the models considered in Section 2.3.3 in the sense that HMMs provide a lot more flexibility in defining the number and characteristics of the regime, while being less interpretable.

2.3.5 Generalized autoregressive conditional heteroskedasticity models

A class of models that is often used to estimate asset return distributions in the context of asset allocation are the Multivariate Generalized Auto-regressive Conditional Heteroskedasticity (MGARCH) models. The main strength of MGARCH models is their ability to simultaneously analyze and forecast the volatility of multiple time series. Carroll et al. (2017) uses MGARCH in combination with a global minimum variance (GMV) optimization strategy to obtain optimal portfolio weights. The reason that Carroll et al. uses GMV optimization as opposed to mean-variance optimization is the fact that MGARCH typically only produces estimates of the volatility of a time series. Using the GMV optimization strategy avoids the problems that are encountered in forecasting expected asset returns, as outlined in Otranto (2010). There are many variations of MGARCH models that could be used to find the optimal portfolio in the minimum variance framework. In addition, Otranto (2010) presents an MGARCH model that incorporates the first moment of the distribution. The author of the study, Otranto (2010), employs the Regime Switching Flexible Dynamic Correlation (RSFDC) model to achieve this. The RSFDC model can be compared with the other models in this section within the framework of mean-variance optimization.

2.4 Asset return distribution model choice

In Section 2.3, various methods for modelling asset return distributions are discussed. The models presented in Section 2.3 aim to capture the stylized facts discussed in Section 2.2 more effectively than the classical model. This overview enables me to answer the first research question from a literature point of view.

1. Which model is the most efficient in capturing stylized effects commonly observed in the return distributions of an institutional investor portfolio over time?

In order to answer the research question, one needs to clearly define "most effective at capturing stylized effects". Hence, it is important to highlight that the following discussion will concentrate on the effects listed in Section 2.2. The initial objective in Section 2.3, was finding models that are effective in capturing the stylized effects often observed in asset returns, while not violating the risk and return MPT constraint. It is already established that a normal distribution in combination with simple historical averages commonly used in MPT, cannot capture fat tails and skewness, two nonlinear effects that are commonly observed in asset return series. A potential solution to this issue lies in the nature of HMMs, as these models excel at capturing skewness and fat tails while also respecting the risk and return constraint of MPT. This is achieved by merging two distinct normal distributions, resulting in a mixture of normal distributions which is not necessarily normally distributed itself and can therefore accommodate more complex patterns in the data. This is illustrated in Figure 2.3.

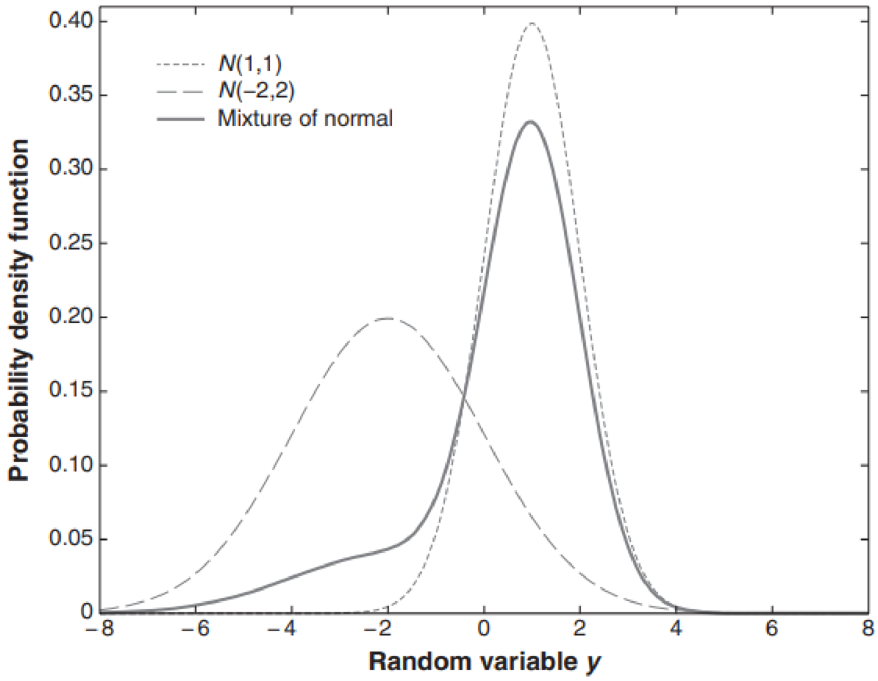

Figure 2.3: Example of combined normal distributions (Ang and Timmermann, 2012).

From looking at Figure 2.3, it is clear that the mixture of two normal distributions is able to produce fat tails and skweness. Therefore, HMMs would potentially be a good fit for modeling asset return series. HMMs are able to effectively capture time-varying effects such as high correlation of assets during market downturns and volatility clustering. This is because HMMs allow for different correlation values between regimes (Ang and Timmermann, 2012). To observe this, please refer to Figure 2.4. Figure 2.4 illustrates the exceedance correlations of US and UK asset returns. The exceedance correlation "estimates the correlation between assets conditional on asset returns falling above or below a pre-specified level" called the exceedance θ (Liu, 2023, p. 3). For an exceedance of $\theta = +2$, the correlation is computed based on observations that exceed three times the mean of the US and UK asset returns. When $\theta = -2$, we compute the correlation based on data that are below -1 times the average in the US and UK asset returns. To see this, let (y_1, y_2) represent two observations drawn from a bivariate variable $Y = (y_1, y_2)$.

Assuming the exceedance level θ is a positive value, we select observations where both y_1 and y_2 exceed θ percent of their respective empirical means. For $\theta \geq 0$, we choose the subset of observations where $y_1 \geq (1+\theta)\bar{y}_1$ and $y_2 \geq (1+\theta)\bar{y}_2$, where $\bar{y}_j$ represents the average value of y_j, and if $\theta \leq 0$, we choose the subset of observations where $y_1 \leq (1+\theta)\bar{y}_1$ and $y_2 \leq (1+\theta)\bar{y}_2$. The correlation of this subset of points is called the exceedance correlation (Ang and Bekaert, 2002). The dashed line corresponds to the exceedance correlations from an HMM, whereas the correlations from the data are shown by squares. The exceedance correlation of a normal distribution and an asymmetric GARCH model are also fitted to the data. From Figure 2.4, it is clear that HMMs significantly outperform all alternative models in terms of fitting exceedance correlations.

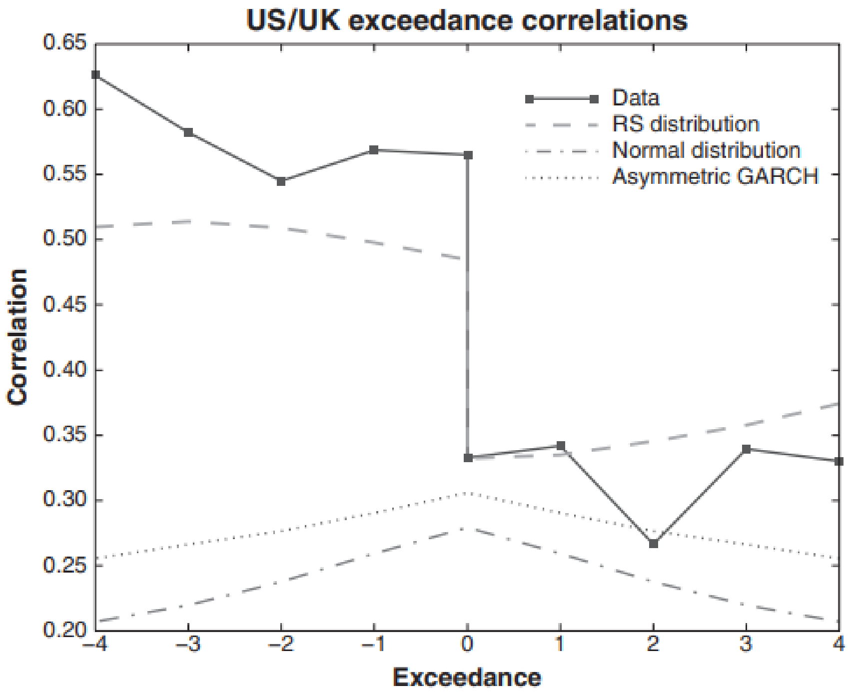

Figure 2.4: Exceedance correlations (Ang and Timmermann, 2012)

Note: The Figure displays exceedance correlations of US and UK asset returns, which are correlations dependent on exceedances θ. Exceedances are expressed as percentages above the empirical mean. For an exceedance of $\theta = +2$, the correlation based on observations that exceed 3 times the mean of the US and UK asset returns is computed. When $\theta = -2$, the correlation based on data that are below -1 times the average in the US and UK asset returns is calculated. The dashed line corresponds to the exceedance correlations from an HMM, whereas the correlations from the data are shown by squares. The exceedance correlation of a normal distribution and an asymmetric GARCH model are also fitted to the data.

The analysis in this section suggests that HMMs perform most effectively in capturing stylized effects in asset return series. HMMs seem to perform better than the alternative models while not violating the risk and return constraints. For this reason, the HMM will be used to model and forecast stock market regimes and asset returns. These forecasts will be used to make new and improved asset allocation decisions.

Chapter 3

Data

This chapter provides a summary of all the possible asset types in the investment universe of an institutional investor and their respective benchmarks. Furthermore, I will present an overview of macroeconomic variables that will be used to analyze stock market patterns within the HMM framework. For both the asset class benchmarks and the macroeconomic variables, the reasoning behind their selection will be provided.

3.1 Data frequency

The literature on hidden Markov models (HMMs) is divided when it comes to the frequency of the data considered. There are several viable options for the data frequency, namely: daily, weekly, and monthly. Daily data was used by Nystrup et al. (2018), weekly data was used by Haase and Neuenkirch (2023), and monthly data was used by Ang and Timmermann (2012); Ang and Bekaert (2002); Guidolin and Timmermann (2007, 2008). Naturally, the data frequency chosen will depend on the objective of the analysis. Initial analysis suggests that a smaller time window would better fit the objective of this research. I refer to Appendix 9.1 for the analysis. Evident from the fact that the statistical accuracy and economic value improve as the data frequency increases. These results can be explained by noticing that it is easier to predict what happens a day from now as opposed to forecasting what will happen a month from today. With this logic, one could assume that the optimal data frequency would be daily.

Considering the practical consequences of this option, it is apparent that this is not feasible. Utilizing daily signals in a tactical asset allocation (TAA) framework would not be logical since daily forecasts cannot be effectively used in a framework where trades are made on a monthly or quarterly basis. The use of daily forecasts would only make sense if the type of asset allocation considered would be market timing (Eychenne et al., 2011). From the paper of Eychenne et al., the conclusion could be made that TAA should use monthly forecasts. However, focusing on monthly data is not logical in terms of the economic value that is generated. A compromise between the two effects is therefore picked, and it is decided to use weekly data.

3.2 Macroeconomic variables

An issue in connecting the business cycle to asset returns through exogenous variables in the context of HMMs is the availability of exogenous variable data Nystrup et al. (2015). Frequently, the data utilized for predictions is unavailable at the time of making a forecast. Consequently, the forecast becomes an in-sample forecast instead of an out-of-sample forecast. Therefore, one should exercise caution when selecting data. The set of macroeconomic variables is selected based on results in the literature and economic common sense.

Hayo et al. (2010); Anzuini et al. (2012); Joyce et al. (2010) all show that central banks, like the Federal Reserve (FED), hold influence over the stock market. The mandate of the FED is to promote price stability and ensure a sustainable level of maximum employment, which is a broader mandate than that of the other major central banks in the developed world, like the European Central Bank, the Bank of England, and the Bank of Japan, which all have the goal of ensuring price stability. The FED has the following tools at its disposal: open market operations, discount window and discount rate, reserve requirements, interest on reserve balances, overnight reverse repurchase agreement facility, term deposit facility and central bank liquidity swaps, among others. The three most relevant tools for monetary policy are open market operations, the discount rate, and reserve requirements (The, 2023). From these three tools, the discount rate, also known as the federal funds rate, is the most interesting for HMM studies.

The discount rate cannot be discussed without first defining the discount window. "The discount window is the process of the Federal Reserve lending to depository institutions and it plays an important role in supporting the liquidity and stability of the banking system and the effective implementation of monetary policy. By providing ready access to funding, the discount window helps depository institutions manage their liquidity risks efficiently and avoid actions that have negative consequences for their customers, such as withdrawing credit during times of market stress" (Fed, 2023a). The discount rate is then defined as "the interest rate charged to commercial banks and other depository institutions on loans they receive from their regional Federal Reserve Bank's lending facility—the discount window" (Fed, 2023a). There are three types of credit offered by the Federal Reserve Banks, namely: primary credit, secondary credit, and seasonal credit. In this study, I only consider the primary credit rate. In the bottom left pane of Figure 3.1 a graphical representation of the primary rate can be found.

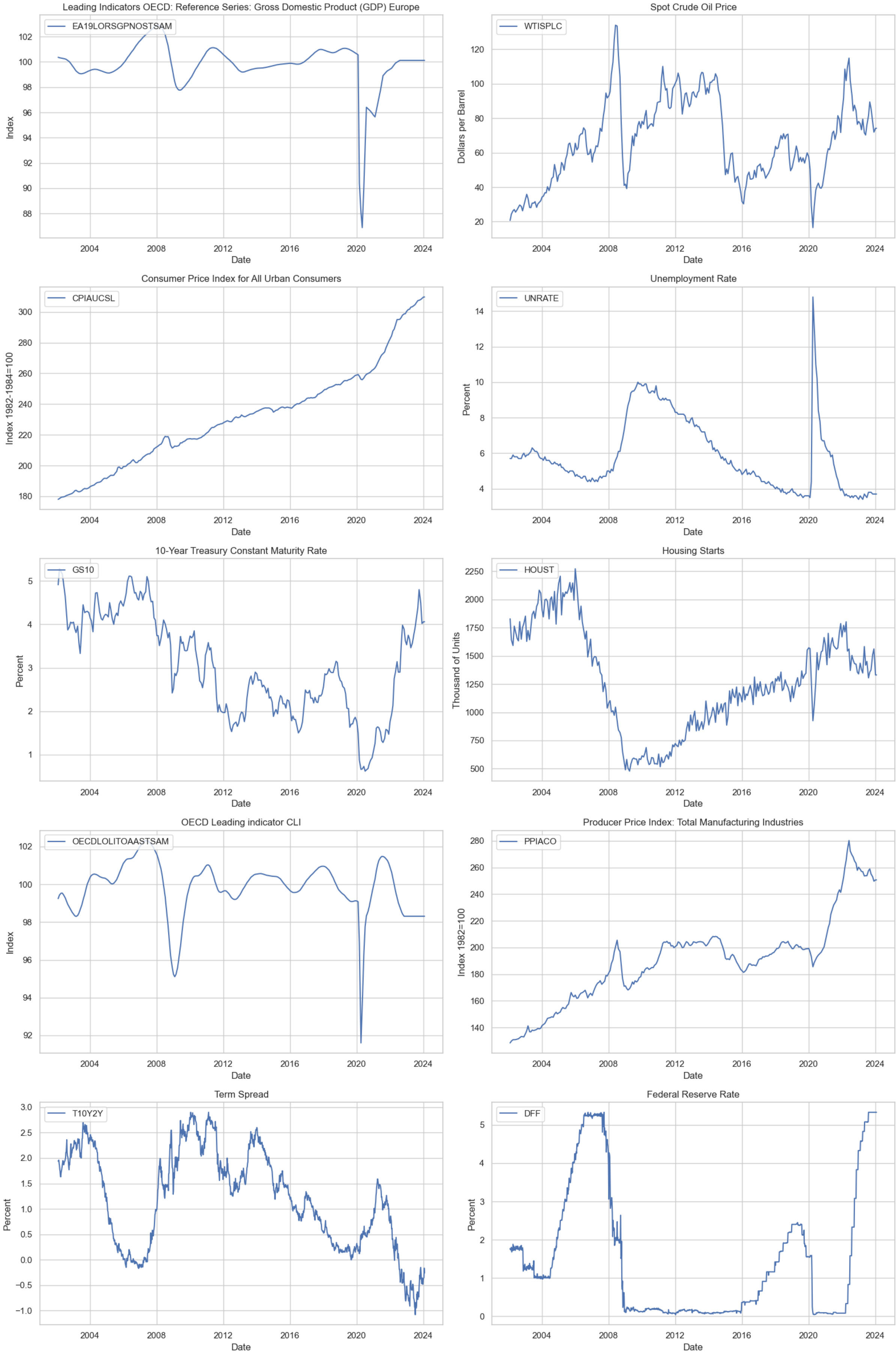

Figure 3.1: Development of macroeconomic variables over the time period from 2002-02-01 to 2024-02-01.

A concept closely related to the federal funds rate is the 10-year Treasury rate. Obviously, interest rate fluctuations have an effect on the cost of borrowing for consumers and (non) listed companies. This can affect their choices to spend and invest, which directly influences the profitability of companies. Changes in interest rate, may thus signal changing economic circumstances that impact stock market returns. Another often-used macroeconomic variable that is related to both the federal funds rate and the 10-year Treasury rate is the term spread. The term spread represents the market expectation for future inflation and economic development. It is calculated as the difference between the 10-year Treasury constant maturity and the 2-year Treasury constant maturity. Stock market returns are impacted by shifts in investor mood and economic outlook, of which the narrowing or widening of the term spread is an often used indicator (Kondlo, 2016; Chen, 2012). The development of these macroeconomic factors is displayed in Figure 3.1. Interestingly, the co-movement between the term spread, the federal funds rate, and the 10-year treasury rate does not seem to be very strong.

It is a well-known fact that commodity prices influence world economic activity and thus stock market returns (Haase and Neuenkirch, 2023). Therefore, the spot crude oil price is included as a macroeconomic variable. It is a good indicator of risk perception, since crude oil prices affect consumer spending, production costs, and inflation. Businesses are impacted by changes in oil prices to various degrees, with energy-dependent sectors likely to be impacted most heavily.

Inflation is another macroeconomic variable that potentially influences the stock market (Haase and Neuenkirch, 2023). A common measure of inflation is the Consumer Price Index (CPI). The CPI acts as a measure of inflation by tracking changes in the price of a basket of products and services. Inflation affects interest rates, consumer buying power, and company profitability. By identifying transitions between inflationary and deflationary regimes, HMMs may provide more insight into variations in risk premia in the stock market. In addition to the CPI, it is also interesting to look at the Producer Price Index (PPI). The PPI tracks the prices domestic producers receive for their goods sold. A rising PPI might be an early signal of rising inflation, which may also have an effect on consumer buying power and business profitability. Switches between inflationary and disinflationary regimes may be detected by using HMMs. Thus providing information on how variations in producer prices might impact stock market returns.

The unemployment rate is a reflection of both the state of the labor market and the nation's overall economic health. High unemployment rates can have negative effects on stock market returns by reducing business profitability and consumer spending (Chen, 2012; Kondlo, 2016). Hence, changes in economic growth and stock market returns may be detected by analyzing information on the dynamics of unemployment. Housing starts is yet another primary measure of economic activity. Housing starts is defined as the number of new privately owned housing units on which construction has started that month. It captures consumer confidence and the desire for new homes. Variations in housing starts have the potential to impact various economic sectors, such as manufacturing, consumer spending, and construction. Because of its wider economic ramifications, the housing market might be a good indicator of the stock market. A perfect example of this is, of course, the 2008 financial crisis, which was in part caused by the housing bubble.

The Organization for Economic Co-operation and Development (OECD) has created a collection of economic indicators known as the OECD Composite Leading Indicators (CLI). The definition of the CLI is "The composite leading indicator (CLI) is designed to provide early signals of turning points in business cycles showing fluctuation of the economic activity around its long term potential level. CLIs show short-term economic movements in qualitative rather than quantitative terms." (Organisation for Economic Co-operation and Development (OECD), 2024) It is often used by economic researchers and policymakers to improve forecasting. The CLI consists of several variables, including interest rate spreads, consumer mood, and business mood. These and other economic factors have been selected based on their track record in predicting major shifts in the economy. The effectiveness of the CLI as a leading indicator can be viewed in Figure 3.2, by noticing that the OECD CLI indicator precedes many of the large economic downturns over the past 30 years. Please note that economic activity is measured by the OECD estimate of economic activity, indicated by a dark orange area if under the long-term trend and a light orange area when above the long-term trend.

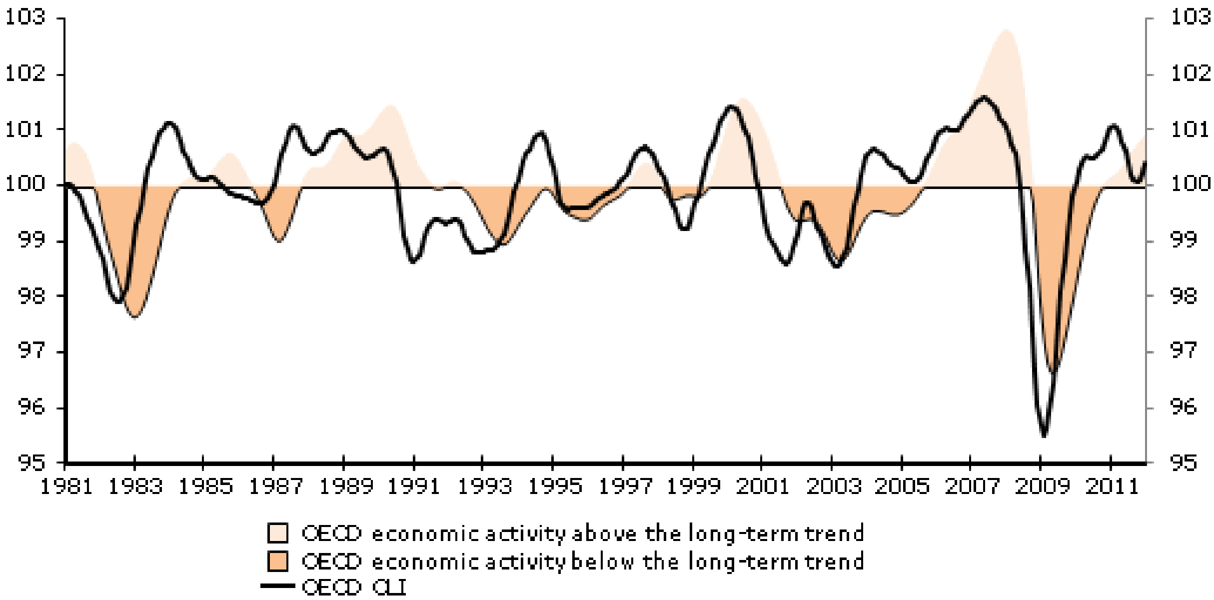

Figure 3.2: Organization for Economic Co-operation and Development (OECD) area composite leading indicator (CLI) compared to the OECD estimate of economic activity, the long-term trend = 100 (Organisation for Economic Co-operation and Development (OECD), 2024).

The total list of macroeconomic factors considered and the way they are accessed can be found in Appendix 9.2.

3.3 Details on asset classes

The type of investors that Sprenkels typically serves, like pension funds and insurers, invest in a broad range of asset classes. Possible investment choices include bonds, stock market indices of either emerging or developed markets, mortgages, real estate, currency, and hedge funds. Some of these asset classes, like real estate, are not publicly traded, meaning that there is no price data available for the analyses in this paper. For these asset classes, the decision is made to look for substitution indices as benchmarks. This method aligns with the current analytical practices at Sprenkels. Sprenkels currently uses publicly available return data supplemented with private data to conduct their analyses. The same data will be used to conduct this research.

These asset classes can be further subdivided into two categories: marketable securities and fixed-income securities. Marketable securities are relatively high-risk in comparison to fixed-income securities. Institutional investors mostly aim to have a healthy mix of 60% marketable securities and 40% fixed income securities in order to keep their risk profile tolerable and within preset bandwidths. Furthermore, it is important to mention that the customers of Sprenkels are exclusively Dutch institutional investors with an interest in the European market.

3.3.1 Marketable securities

Stock index

For the developed market stock index, the MSCI World is chosen as a representative index. The MSCI World is a commonly selected index in the literature (Nystrup et al., 2018). For the emerging market stock index, the MSCI emerging market (EM) index is chosen as a representative index; it is also often found in the portfolios of institutional investors.

High-Yield Corporate debt

High-yield corporate debt, otherwise known as "junk bonds," is often considered an interesting investment class. High-yield corporate bonds are typically issued by companies with lower credit ratings; a lower credit rating often means they are rated below BBB by major credit rating agencies such as Standard & Poor's. These lower credit ratings reflect a higher risk of default on the part of the issuing company. High-yield corporate debt is an interesting alternative to stock indices. Furthermore, emerging market (EM) high-yield corporate bonds are also considered.

Commodities

For commodities, two specific asset classes have been chosen, namely gas and gold. These asset classes have been chosen because they are the asset classes that are historically often selected by investors. However, interest in these commodities has been steadily declining, according to the experts at Sprenkels.

Real estate

When discussing non-listed real estate, it is important to realize that "dynamic asset allocation is, by definition, more restricted than SAA in terms of the size of the investment opportunity set because it is difficult to invest dynamically in illiquid assets such as private real estate, private equity, infrastructure, timber, etc. This is worth mentioning, given that illiquid alternatives have become a larger part of institutional investors' portfolios in recent years." (Nystrup et al., 2018, p .1) This implies that illiquid asset classes might be less suited for HMM-based models, something that should be taken into account. For now, real estate will be included in the investment universe.

3.3.2 Fixed-income securities

Government Bonds

It is also necessary to define assets that carry lower levels of risk. Firstly, the 13-week Treasury bill, often referred to as the "three month Treasury bill" or "T-bill." Secondly, the portfolio of an institutional investor will also contain long-term bonds. For the long-term maturity bonds, 20-year maturity treasury bonds will be utilized.

3.3.3 Overview of asset classes

Data on all of these assets is imported and transformed into the correct format. Below, in Figure 3.3, the returns on all of these assets can be found. One can notice that the MSCI World index and the real estate index far outperform all of the other asset classes in terms of cumulative wealth but seem highly correlated. Furthermore, periods of downturn like the COVID-19 recession hit almost all of the asset classes equally hard, except for fixed-income securities.

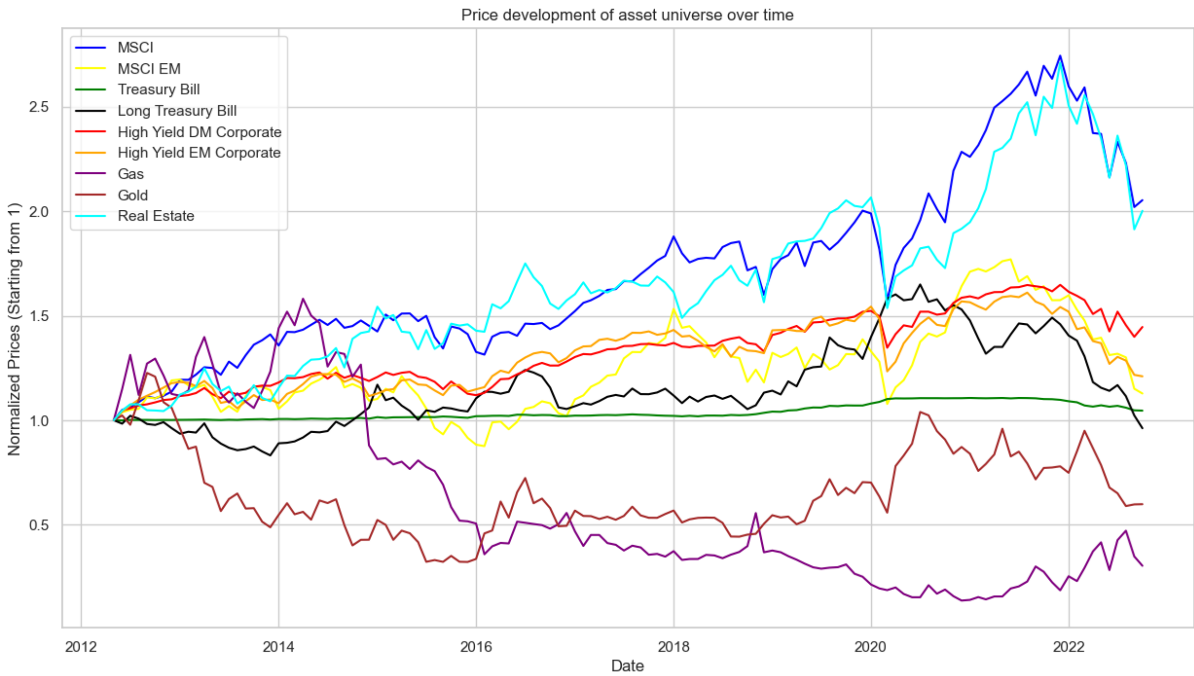

Figure 3.3: Development of selected asset classes over time period from 2012-04-01 to 2024-02-01.

Additional information on all of these assets can be found in Table 3.1. As mentioned, the MSCI World and the real estate index yield higher returns than the other asset classes, but this comes at a higher level of volatility. Please note that the MSCI EM index has a higher volatility than the MSCI while yielding lower returns. Furthermore, it is not strange to see that many institutional investors lost interest in commodities over the past 10 years when looking at the price development in Figure 3.3.
Furthermore, it is also interesting to look at the correlations between the various asset classes to further study the degree of their co-movements. Table 3.2 gives an overview of all the different asset classes

	Mean Return (%)	Volatility (%)	Skewness	Kurtosis
MSCI	0.170	2.145	-0.550	6.437
MSCI EM	0.067	2.636	-0.233	1.396
T-bill	0.016	0.177	0.788	7.487
20+ Bond	0.045	1.964	-0.275	0.859
DMHYC	0.088	1.352	0.371	30.681
EMHYC	0.072	1.479	-2.061	21.511
Gas	-0.202	6.447	0.023	0.983
Gold	0.079	4.965	-0.398	4.846
Real estate	0.176	2.952	-0.210	14.730

Table 3.1: Summary statistics of selected asset classes.

and their correlations. The observation can be made that assets like the MSCI World and high-yield corporate bonds in both emerging markets and developed markets are highly correlated. While alternative investments like gas, gold, and Treasury bonds are not. Notably, real estate has a high level of correlation with the stock market.

	MSCI	MSCI EM	T-bill	20+ Bond	DMHYC	EMHYC	Gas	Gold	Real estate
MSCI	1.00	0.76	-0.04	-0.12	0.72	0.80	0.13	0.23	0.70
MSCI EM	0.76	1.00	0.07	-0.06	0.77	0.67	0.11	0.30	0.50
Tbill	-0.04	0.07	1.00	0.67	0.08	0.16	0.00	0.15	0.15
20+ Bond	-0.12	-0.06	0.67	1.00	0.07	0.08	-0.10	0.20	0.28
DMHYC	0.72	0.77	0.08	0.07	1.00	0.78	0.11	0.34	0.62
EMHYC	0.80	0.67	0.16	0.08	0.78	1.00	0.17	0.34	0.69
Gas	0.13	0.11	-0.00	-0.10	0.11	0.17	1.00	-0.05	0.15
Gold	0.23	0.30	0.15	0.20	0.34	0.34	-0.05	1.00	0.23
Real estate	0.70	0.50	0.15	0.28	0.62	0.69	0.15	0.23	1.00

Table 3.2: Correlation Matrix of selected asset classes.

From the analysis of Figure 3.3, Table 3.1, and Table 3.2, it becomes apparent that the current selection of assets is a diverse representation of the many investment opportunities available to an institutional investor.

Chapter 4

Methodology

The following sections contain a detailed overview of the hidden Markov models (HMM). First, I will discuss the overall framework of the model, whereafter the process of obtaining estimates will be outlined. Then, further information on forecasting will be provided.

4.1 Hidden Markov model

In order to deal with the non-linear nature of asset returns, a particular set of regime-based asset allocation (RBAA) models called the k-regime hidden Markov models (HMMs) has been selected for the modeling process. The first application of HMMs can be traced back to 1989, when Hamilton (1989) in his seminal work applied these models to the business cycle, identifying recessionary and expansionary periods in time. However, these HMMs have also been demonstrated to be effective in modeling time series of asset returns, playing a valuable role in the asset allocation decision-making process, as shown by Ang and Bekaert (2002); Guidolin and Timmermann (2007); Nystrup et al. (2018); Haase and Neuenkirch (2023). The canonical model used in regime switching analysis is of the following form: (Ang and Timmermann, 2012; Hamilton, 2010, 2020):

$$y_t = \mu_{s_t} + \phi_{s_t} y_{t-1} + \sigma_{s_t} \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, 1). \tag{4.1}$$

Upon closer inspection, one can observe that this notation is very similar to that of a vector autoregression model of the first order VAR(1). In this paper, y_t represents the returns on various asset classes over time. The exact asset classes that are considered in this paper can be found in Chapter 3. Naturally, asset returns y_t could be dependent on their own past values y_{t-1}, the level of correlation is dictated by the autocorrelation ϕ_{s_t}. In addition, the asset return process is affected by random shocks ε_t, which are assumed to be independent and identically distributed (i.i.d.) with mean 0 and standard deviation σ_{s_t}. Lastly, the model has an intercept μ_{s_t}. The notation of the canonical HMM differs from an VAR(1) model in the fact that the intercept μ_{s_t}, standard deviation σ_{s_t}, and autocorrelation ϕ_{s_t} are all affected by a discrete random variable $s_t \in (0, 1, 2, \ldots, k)$, where k is equal to the number of regimes the process can possibly visit. The discrete random variable s_t is the so-called regime variable; it governs the process of regime switching in our model. Commonly, the variable s_t follows a first-order homogeneous Markov chain, hence the name hidden Markov model (HMM). For a first-order homogeneous Markov chain, the transition probabilities between regimes are given by the following equation:

$$\Pr(s_t = j \mid s_{t-1} = i) = p_{ij}, \quad i, j \in (0, 1, 2, \ldots, k). \tag{4.2}$$

One of the attractive features of the HMM is its relative simplicity, evident from the fact that the model is fully defined with the help of equations (4.1) and (4.2). An useful extension of the canonical model in equation (4.1) would be to allow for our dependent variable stock market returns to be regressed on exogenous variables like, for example, the dividend yield. The model extension is justified by Kim et al. (1999):

$$y_t = \mu_{s_t} + \phi_{s_t} y_{t-1} + \beta x_{t-1} + \sigma_{s_t} \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, 1). \tag{4.3}$$

The derivation of properties for the model in equation (4.1) and the model in equation (4.3) follow the same steps. For ease of notation, the derivations in the rest of this chapter will only be done for the model in equation (4.1); calculations for the model in equation (4.3) follow analogously.

4.2 Estimation

In order to estimate the model parameters, it is imperative to outline a procedure tailored for this purpose. This procedure will be derived in accordance with the canonical model outlined in Section 4.1. It is important to note that the subsequent derivation assumes the validity of certain properties that are defined below. These assumptions are made, partly based on the resulting computational convenience:

- Time homogeneous transition probabilities.
- One asset.
- Two states, a bull and a bear state.
- Markov chain is of the first order.
- The random shocks are Gaussian.

Please note that the derivations outlined in the rest of this chapter can easily be extended to more sophisticated models beyond the baseline model. Now that the model is fully defined, the next step will be to define the estimation procedure. Firstly, note that the value of the regime variable s_t is not directly observed, which is a natural consequence of the fact that in the stock market, a trader is also never certain whether the market is in a bull market or a bear market. His best bet is to look at the performance of the stocks making up the market and make a guess based on patterns observed. More formally, it is possible to obtain a probabilistic inference about the value of the regime variable s_t by observing the asset return process y_t (Hamilton, 2010). The fact that the regime process s_t is unobserved is the reason that these models are often also described as hidden Markov models (HMMs). The initial set of parameters that we want to infer are the state probabilities ξ_{jt}, which are defined as:

$$\xi_{jt} = \Pr\left(s_t = j \mid \Omega_t; \theta\right), \quad j = 1, 2, \tag{4.4}$$

where:

- $\sum_{j=1}^{2} \xi_{jt} = 1$.
- $\boldsymbol{\theta} = (p_{00}, p_{10}, p_{01}, p_{11}, \mu_1, \mu_2, \sigma_1, \sigma_2, \phi_1, \phi_2)'$.
- $\Omega_t = (y_t, y_{t-1}, y_{t-2}, ..., y_1, y_0)'$.

For now, we assume $\boldsymbol{\theta}$ to be known. The inference process is iterative in nature, meaning that the procedure will start at $t = 0$ with ξ_{j0} and will iterate forward for $t = 1, 2, \ldots, T$. Before continuing the derivation, some important quantities will need to be defined.

$$\xi_{i,t-1} = \Pr\left(s_{t-1} = i \mid \Omega_{t-1}; \boldsymbol{\theta}\right), \quad i = 1, 2. \tag{4.5}$$

With the conditional probability of being in a certain state defined, the next step is to define the conditional probability density of y_t. The conditional probability density function is given by equation (4.6). This formula follows rather trivially from the assumption that the random shocks ε_t follow a Gaussian distribution.

$$\eta_{jt} = f\left(y_t \mid s_t = j, \Omega_{t-1}; \boldsymbol{\theta}\right) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{\left(y_t - \mu_j - \phi_j y_{t-1}\right)^2}{2{\sigma_j}^2}\right]. \tag{4.6}$$

Moreover, the conditional probability density defined in equation (4.6) can be extended to the conditional density function (CDF) of y_t given Ω_{t-1}, $\boldsymbol{\theta}$, and without conditioning on the regime s_t. To observe this extension, it is evident that η_{jt} and $f\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right)$ are akin expressions, differing only in the former's conditioning on the state variable s_t. By multiplying $\xi_{i,t-1}$, the probability of being in state i at time $t-1$, with the transition probabilities p_{ij}, we obtain the probabilities of being in state j at time t. When this is further multiplied by η_{jt}, it results in the conditional density $f\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right)$. Formally:

$$f\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right) = \sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij} \xi_{i,t-1} \eta_{jt}. \tag{4.7}$$

Now that the conditional density is determined, it becomes possible to find an expression for ξ_{jt}. This expression is obtained by using the 'Bayes-rule' and some elementary probability theory.

$$
\begin{aligned}
\xi_{jt} = \Pr\left(s_t = j \mid \Omega_t; \boldsymbol{\theta}\right) &= \frac{\Pr\left(y_t, \Omega_{t-1}, s_t = j; \boldsymbol{\theta}\right)}{\Pr\left(y_t, \Omega_{t-1}; \boldsymbol{\theta}\right)} \\
&= \frac{\Pr\left(y_t \mid s_t = j, \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(s_t = j, \Omega_{t-1}; \boldsymbol{\theta}\right)}{\Pr\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(\Omega_{t-1}; \boldsymbol{\theta}\right)} \\
&= \frac{\Pr\left(y_t \mid s_t = j, \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(s_t = j \mid \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(\Omega_{t-1}; \boldsymbol{\theta}\right)}{\Pr\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(\Omega_{t-1}; \boldsymbol{\theta}\right)} \\
&= \frac{\Pr\left(y_t \mid s_t = j, \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(s_t = j \mid \Omega_{t-1}; \boldsymbol{\theta}\right)}{\Pr\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right)}
\end{aligned}
\tag{4.8}
$$

Substituting the expression for η_{jt} defined in equation (4.6) and the conditional density defined in equation (4.7) into equation (4.8) yields the following:

$$
\begin{aligned}
\frac{\Pr\left(y_t \mid s_t = j, \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(s_t = j \mid \Omega_{t-1}; \boldsymbol{\theta}\right)}{\Pr\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right)} &= \frac{\eta_{jt} \cdot \Pr\left(s_t = j \mid \Omega_{t-1}; \boldsymbol{\theta}\right)}{\Pr\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right)} \\
&= \frac{\eta_{jt} \cdot \Pr\left(s_t = j \mid \Omega_{t-1}; \boldsymbol{\theta}\right)}{\sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij} \xi_{i,t-1} \eta_{jt}}.
\end{aligned}
\tag{4.9}
$$

Please note that $\Pr\left(s_t = j \mid \Omega_{t-1}; \boldsymbol{\theta}\right)$ in equation (4.9) can be rewritten in the following way:

$$
\Pr\left(s_t = j \mid \Omega_{t-1}; \boldsymbol{\theta}\right) = \sum_{i=1}^{2} \Pr\left(s_{t-1} = i \mid \Omega_{t-1}; \boldsymbol{\theta}\right) \cdot \Pr\left(s_t = i \mid s_{t-1} = j\right),
$$

which, if substituted into equation (4.9) yields an expression of η_{jt}:

$$
\xi_{jt} = \Pr\left(s_t = j \mid \Omega_t; \boldsymbol{\theta}\right) = \frac{\sum_{i=1}^{2} p_{ij} \xi_{i,t-1} \eta_{jt}}{\sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij} \xi_{i,t-1} \eta_{jt}}. \tag{4.10}
$$

With ξ_{jt} defined, it becomes possible to perform inference of the probabilities ξ_{jt} for $t = 1, 2, \ldots, T$. By starting from an initial probability ξ_{i0} and iterating forward with the help of equation (4.10). In the process, the sample conditional log likelihood of the observed data defined in equation (4.11) will also be obtained.

$$
\log f\left(y_1, y_2, \ldots, y_T \mid y_0; \boldsymbol{\theta}\right) = \sum_{t=1}^{T} \log f\left(y_t \mid \Omega_{t-1}; \boldsymbol{\theta}\right). \tag{4.11}
$$

The algorithm evaluates the sample log likelihood based on a specified value of $\boldsymbol{\theta}$ that was set before starting the iteration. However, with the log likelihood function finally obtained, it is now possible to obtain estimates for the value of $\boldsymbol{\theta}$ with the help of numerical optimization methods. The procedure outlined above is called the Hamiltonian filter, and it was first developed by Hamilton (1989).

4.3 Initialization

Before commencing the iteration process, it is essential to initialize the model. This involves specifying the starting values, denoted as ξ_{i0}, of the algorithm. In case the Markov chain is ergodic, implying that it eventually visits every possible state, the unconditional probability version of equation (4.4) can be used to obtain initial probabilities, in the simple two-state canonical model, this leads to the following two starting probabilities: (Hamilton, 2010, 2020):

$$\begin{aligned} \xi_{10} &= \frac{1 - p_{22}}{2 - p_{11} - p_{22}}. \\ \xi_{20} &= \frac{1 - p_{11}}{2 - p_{22} - p_{11}}. \end{aligned} \tag{4.12}$$

These equations calculate the probability of the initial state being $i = 1, 2$ based on the ergodicity assumption, where p_{11} and p_{22} represent the transition probabilities. In addition, the number of possible regimes k needs to be specified. Deducing the number of regimes from the data is generally difficult, so it is advised to base the choice on economic arguments (Ang and Timmermann, 2012). From an economic point of view, it seems logical to set the number of possible regimes equal to two, stemming from the fact that the theory of there being a so-called bear state and a so-called bull state is widely supported in the financial world. This choice is further supported by the literature (Ang and Bekaert, 2002; Haase and Neuenkirch, 2023).

4.4 Smoothed inference

As a result of using the Hamilton filter, one obtains the so-called filtered probabilities. These can be quite rugged, which means that they might not be suitable for forecasting or interpretation. One of the common ways to deal with this is smoothing; this can be done with Kim's smoother (Kim, 1994). Smoothed probabilities are defined in the following way:

$$\xi_{t|\tau} = \Pr\left(s_t = i \mid \Omega_\tau; \boldsymbol{\theta}\right), \quad i = 1, 2, \tag{4.13}$$

where $t < \tau$. The basic idea behind Kim's smoother is that one uses all of the information in the sample to create an approximation that removes excess noise from the results. We will not go into detail since smoothing is done with the help of packages in Python.

4.5 Obtaining parameter estimates

As mentioned in Section 4.2, the log-likelihood function as defined in equation (4.11) needs to be maximized in order to find the unknown parameters $\boldsymbol{\theta}$. A commonly used technique for maximizing the maximum likelihood function is the expectation maximization (EM) algorithm (Hamilton, 1989, 2016). The EM algorithm procedure generates a sequence of $\boldsymbol{\theta}^{(i)}$, one for each step i in the iteration. The EM algorithm ensures that a local maximum is attained by checking at every iteration that the log-likelihood function in equation (4.11) of $\boldsymbol{\theta}^{(i+1)}$ is higher than the log-likelihood of $\boldsymbol{\theta}^{(i)}$. As stated, the starting $\boldsymbol{\theta}$ is often chosen based on an educated guess. In order to obtain estimates of $\boldsymbol{\theta}^{(i+1)}$ at every step, we use the estimate of $\boldsymbol{\theta}^{(i)}$ to evaluate the smoothed probabilities $xi_{t|\tau}$ as defined in equation (4.13). These smoothed probabilities are then used to find the new transition probabilities and new parameters $\boldsymbol{\theta}^{(i+1)}$. The new estimate of $\boldsymbol{\theta}^{(i+1)}$ is then weighted by the corresponding smoothed probability, and these weighted parameter values are then used to calculate the log-likelihood function. As said, the algorithm stops if, at some point, we fail to obtain a higher log-likelihood value, ensuring that a local maximum is attained. The exact math can be found in Hamilton (2016). A potential pitfall of this modeling approach is the fact that there could be multiple local maxima. As a remedy, Hamilton (2016) suggests starting the iterations from a large number of different starting points in order to ensure that the algorithm always ends up at one point. The exact methods used to implement the techniques in this section can be found in Appendix 9.3.

4.6 Forecasting

After fitting a model to the data, the next step is to try and forecast future regimes and future stock returns. In order to be able to forecast, a procedure needs to be defined that can be used to forecast the entities of interest. Forecasting can be divided into two categories: forecasting unobserved variables and forecasting observed variables.

4.6.1 Forecasting unobserved variables

Forecasting the unobserved variables, which in the application of this paper are the regimes prevailing in the stock market, is done by following the procedure outlined in the book of Hamilton (2020). The optimal one-period-ahead forecast is given by the following formula:

$$\xi_{t+1|t} = \Pr\left\{s_{t+1} = j \mid \Omega_t; \boldsymbol{\theta}\right\} = \sum_{i=1}^{2} p_{ij} \cdot \xi_{t|t} \quad j = 1, 2, \tag{4.14}$$

where the probabilities p_{ij} correspond to equation (4.2) and where $\xi_{t|t}$ is defined as in equation (4.13).

4.6.2 Forecasting observed variables

The proof below only concerns the one-period forecast but can be analogously extended to multiple periods. Once again, the derivation below follows the assumptions of the model that can be viewed in equation (4.1). The forecast y_{t+1} conditional on Ω_t, s_{t-1}, and $\boldsymbol{\theta}$ is easily found by using equation (4.6). The following expression thus holds for the forecast:

$$E\left(y_{t+1} \mid s_{t+1} = j, \Omega_t; \boldsymbol{\theta}\right) = \mu_j + \phi_j y_t. \tag{4.15}$$

With the help of the expression in equation (4.15), the expression for the forecast of y_{t+1} conditional on Ω_t and $\boldsymbol{\theta}$ can be obtained.

$$
\begin{aligned}
E\left(y_{t+1} \mid \Omega_t; \boldsymbol{\theta}\right) &= \int y_{t+1} \cdot f\left(y_{t+1} \mid \Omega_t; \boldsymbol{\theta}\right) dy_{t+1} \\
&= \int y_{t+1} \left\{ \sum_{j=1}^{2} \Pr\left(y_{t+1}, s_{t+1} = j \mid \Omega_t; \boldsymbol{\theta}\right) \right\} dy_{t+1} \\
&= \int \mathbf{y}_{t+1} \left\{ \sum_{j=1}^{2} \left[f\left(y_{t+1} \mid s_{t+1} = j, \Omega_t; \boldsymbol{\theta}\right) \Pr\left\{ s_{t+1} = j \mid \Omega_t; \boldsymbol{\theta} \right\} \right] \right\} dy_{t+1} \\
&= \sum_{j=1}^{2} \Pr\left\{ s_{t+1} = j \mid \Omega_t; \boldsymbol{\theta} \right\} \int y_{t+1} \cdot f\left(y_{t+1} \mid s_{t+1} = j, \Omega_t; \boldsymbol{\theta}\right) dy_{t+1} \\
&= \sum_{j=1}^{2} \Pr\left\{ s_{t+1} = j \mid \Omega_t; \boldsymbol{\theta} \right\} \cdot E\left(y_{t+1} \mid s_{t+1} = j, \Omega_t; \boldsymbol{\theta}\right).
\end{aligned}
\tag{4.16}
$$

The equation (4.16) can now be used to obtain our forecasts. The forecasting procedure that will be employed in this paper only requires one-step forecasts, which means the above derivation is sufficient.

4.7 Sojourn time distribution

The sojourn time distribution is another significant assumption that the user of HMMs implicitly make. The sojourn time in the context of HMMs is defined as the amount of time a system or process spends in a particular state before transitioning to another state Nystrup et al. (2018). In the context of this paper, the sojourn time distribution is assumed to be geometrically distributed:

$$
\Pr\left(\text{'staying } t \text{ time steps in state } i\text{'} \right) = \gamma_{ii}^{t-1} \left(1 - \gamma_{ii}\right). \tag{4.17}
$$

The implicit assumption of a geometric distribution directly implies that the transition probabilities are independent of the time spent in a certain regime. This is, of course, a logical consequence of the assumption that the Markov chain is modeled as a first-order Markov chain. Alternative distributions have been explored in the literature, for example, in Bulla and Bulla (2006); Guédon (2003). The introduction of these alternative distributions and the subsequent drop of the memory-less property complicate the computations significantly, so much so that the decision is made that these will not be further explored in this paper. However, exploring these alternative distributions is an interesting avenue for further research.

4.8 Time-varying transition probability

Up until now, the HMMs have been of the type time-constant transition probabilities (TCTP), meaning that the probabilities of transitioning from one state to another as defined in equation (4.2) will remain constant over time. The assumption of TCTP might not be correct, since there are very compelling economic arguments to be made for the fact that economic circumstances can change structurally over time. Diebold et al. (1993) in their paper give an example of exchange rate revaluations, of which the likelihood increases under progressively more overvaluation or undervaluation. Such situations could also occur in the stock market. In order to account for these changing circumstances, one can introduce the concept of time-varying transition probabilities (TVTP). Introducing TVTP means that equation (4.2) is altered to the following equation (Diebold et al., 1993; Haase and Neuenkirch, 2023):

$$\begin{aligned}
p_t^{00} &= P\left(s_t = 0 \mid s_{t-1} = 0, x_{t-1}; \beta_0\right) = \frac{\exp\left(x_{t-1}'\beta_0\right)}{1+\exp\left(x_{t-1}'\beta_0\right)} \\
p_t^{01} &= \left(1 - p_t^{00}\right) = P\left(s_t = 1 \mid s_{t-1} = 0, x_{t-1}; \beta_0\right) = 1 - \frac{\exp\left(x_{t-1}'\beta_0\right)}{1+\exp\left(x_{t-1}'\beta_0\right)} \\
p_t^{10} &= \left(1 - p_t^{11}\right) = P\left(s_t = 0 \mid s_{t-1} = 1, x_{t-1}; \beta_1\right) = 1 - \frac{\exp\left(x_{t-1}'\beta_1\right)}{1+\exp\left(x_{t-1}'\beta_1\right)} \\
p_t^{11} &= P\left(s_t = 1 \mid s_{t-1} = 1, x_{t-1}; \beta_1\right) = \frac{\exp\left(x_{t-1}'\beta_1\right)}{1+\exp\left(x_{t-1}'\beta_1\right)},
\end{aligned} \tag{4.18}$$

where $x_{t-1} = \left(1, x_{1,t-1}, \ldots, x_{(b-1),t-1}\right)'$ and $\beta_i = \left(\beta_{i0}, \beta_{i1}, \ldots, \beta_{i(b-1)}\right)', i = 1, 2.$

Here, s_t should once again be understood as the state variable governing the regime-switching process. The two transition probabilities are now modeled as logistics functions and are changing over time. Here, b represents the number of structural breaks in the transition probability process. The transition probability proces is being governed by a set of exogenous variables. The exogenous variables are typically the same variables that enter the model in equation (4.3) as the exogenous variables; however, this does not need to be the case (Diebold et al., 1993; Kim, 1994). In order to relate the TVTP equations to the TCTP equations, one easily verifies that by setting the last $(b-1)$ terms of β_i to 0, one obtains the standard TCTP model.

The derivation of the state probabilities and the log-likelihood functions was first performed by Diebold et al. (1993); it mostly uses the same steps outlined in Section 4.2. Furthermore, obtaining parameter estimates will again be done with the help of the EM algorithm.

4.9 Dimension reduction techniques

The more macroeconomic variables that are included in an HMM, the higher the model uncertainty will be. In order to include a lot of variables but avoid overfitting, one can use dimension reduction methods like principal component analysis (PCA). This paper will follow the work of Haase and Neuenkirch (2023) in the creation and use of PCA in the context of asset allocation models. The goal of PCA is to reduce the dimensionality of a dataset by transforming the original variables into a new set of variables called principal components. Every principal component aims to capture the maximum amount of variability in the data. With the help of a PCA analysis, the set of macroeconomic variables considered is reduced to a set of principal components. In Figure 4.1, the scree plot belonging to the aforementioned PCA analysis can be found.

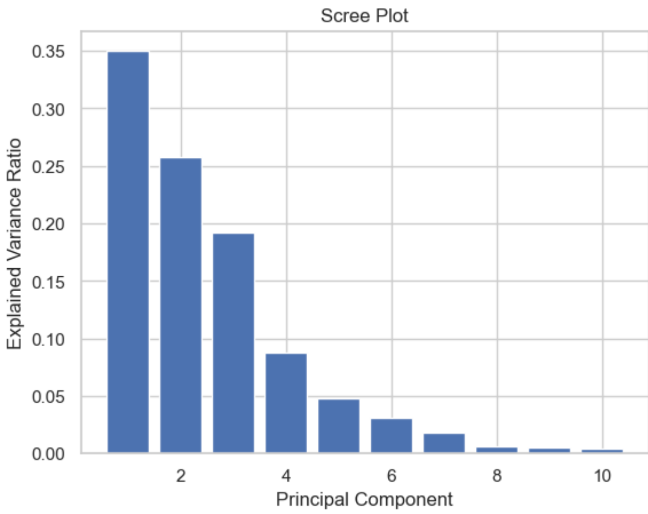

Figure 4.1: Scree plot of PCA analysis on macroeconomic variables

Note: the scree plot provides an overview of the proportion of the overall variance in the macroeconomic variables that is accounted for by each of the principal components.

The scree plot provides an overview of the proportion of variance that the principal components explain. From Figure 4.1, it is evident that the first two principal components explain more than 60% of the variance of the original dataset, suggesting that including these two principal components should be sufficient.

Chapter 5

Results

This paper aims to find the best possible forecast of future regimes and returns, given all of the data available at that specific point in time. Naturally, in order to be able to forecast, a model is required to not only fit the training data but also be able to generalize well to unseen data. However, there is an implicit trade-off made between finding a good forecast and finding a good estimate of the effect parameter that one might be interested in, which is called inference (James et al., 2023). The trade-off between the two has to do with the fact that certain more elaborate and restrictive models might lead to better inferences of relationships between variables in the sample, while less restrictiveness in general leads to better forecasting results. This effect is called overfitting, which means that a model does not generalize well to new data. This negatively impacts the forecasting ability of the model (James et al., 2023). Circling back to the aim of this paper, namely finding good forecasts of futures regimes and returns, one would expect that the models considered in the rest of this chapter would all be very nonrestrictive in nature. However, this is not the case, since a certain level of inference and model interpretability is required in the context of this paper since pension funds and insurers cannot work with black-box solutions.

5.1 Fitting the model

In order to be able to evaluate the fit of the model, appropriate model selection criteria need to be chosen. The first criterion that is a natural choice to consider when starting the model selection process in the context of this paper is the log-likelihood criterion. The log-likelihood criterion follows equation (4.11). The log-likelihood function is calculated as a byproduct of the iteration process that was defined to calculate the state probabilities. However, as previously stated, it is important that the forecasting models in this paper perform well on the trade-off between model fit and model complexity. A criterion like the log-likelihood is not able to make this trade-off since it only seeks to maximize the fit of the model, irrespective of the number of parameters selected. The Bayesian Information Criterion (BIC), defined in equation (5.1), is able to make this trade-off by penalizing complex models with many parameters while considering the fit of these models by looking at the log-likelihood function.

$$\textbf{Bayesian Information Criterion (BIC)} = k \cdot log(N) - 2 \cdot log(\hat{L}) \tag{5.1}$$

A selection of hidden Markov models (HMMs) is chosen based on an examination of the literature, preliminary tests, and economic intuition. The chosen models can be categorized into two categories, namely, the time-constant transition probabilities (TCTP) models and the time-varying transition probabilities (TVTP). TCTP models are predicated on the assumption of constant probabilities of switching over time. These models, popularized by Ang and Bekaert (2002); Nystrup et al. (2018), offer a simplified framework compared to their TVTP counterparts. However, a subset of TVTP models is also explored, inspired by the works of Haase and Neuenkirch (2023); Diebold et al. (1993). These TVTP models allow for more flexibility in capturing regime-switching dynamics, which can be particularly relevant given the complexities and time-varying nature of the financial markets.

This study does not follow one particular strand of literature, as it draws insights from various sources, in particular Ang and Timmermann (2012); Haase and Neuenkirch (2023); Nystrup et al. (2018). Comparing outcomes across different studies in order to find the academic consensus on the most effective models proved challenging due to differences in estimation methodologies stemming from model assumptions (e.g., sojourn time distribution), data frequency, and research objectives. Given the eclectic approach of this study, it is not viable to rely on the conclusions of any single study. Furthermore, it is essential to emphasize that while the selected models offer a robust framework for this study, they do not encompass all possible variations. Nevertheless, they represent a comprehensive cross-section that fits the objectives of this paper. Table 5.1 contains 21 different model specifications. For the constant variance model, only the intercept (μ) is regime-dependent, while for the switching variance models, the intercept (μ) and the random shocks (Ω) are regime-dependent. The autoregressive order specifies the number of lags that are taken into account, while the exogenous variable specifies the number and type of macroeconomic or PCA-generated variables that are included.

All of the 21 model specifications defined in Table 5.1 are fitted to a dataset comprising weekly returns on the MSCI World from 1987-01-01 to 2024-02-01). The model fit of these model specifications is evaluated by looking at the BIC score and the log-likelihood. The results are summarized in Table 5.2 The use of an extensive dataset tries to ensure that the chosen models are well-suited to address the intricacies of asset allocation within the context of stock market dynamics.

It is evident from Table 5.2 that the best scoring model specifications based on the log-likelihood and BIC criteria are those belonging to the TCTP class. Within the TCTP class of models, there are insignificant differences in terms of log-likelihood scores. However, when looking at the BIC scores, it is clear that the simpler models outperform the more elaborated models, which is a direct consequence of the similarity in log-likelihood scores between models. Model specification 2 performs best on both the model selection criteria, meaning that it has the best fit and is best suited for inference. However, as stated, a good model fit does not guarantee the best results in terms of forecasting.

Table 5.1: Overview of HMM specifications

Model	HMM Type	AR order	Type of Variance	Exogenous variable
1	TCTP	0	Constant variance	
2	TCTP	0	Switching variance	
3	TCTP	1	Switching variance	
4	TCTP	1	Switching variance	Term spread
5	TCTP	1	Switching variance	Federal funds rate
6	TCTP	1	Switching variance	Non standardized Term spread
7	TCTP	1	Switching variance	Non standardized federal funds rate
8	TCTP	0	Switching variance	Term spread
9	TCTP	0	Switching variance	Federal funds rate
10	TCTP	0	Switching variance	OECD CLI
11	TCTP	1	Switching variance	One principal component
12	TCTP	1	Switching variance	Two principal components
13	TCTP	1	Switching variance	OECD CLI
14	TCTP	1	Switching variance	Manufacturing Industries
15	TVTP	0	Switching variance	
16	TVTP	1	Switching variance	
17	TVTP	1	Switching variance	Term spread
18	TVTP	1	Switching variance	Federal funds rate
19	TVTP	1	Switching variance	OECD CLI
20	TVTP	1	Switching variance	One principal component
21	TVTP	1	Switching variance	Two principal components

Notes: This table contains 21 different model specifications. For the constant variance model only the intercept (μ) is regime dependent, while for the switching variance models the intercept (μ) and the random shocks (Ω) are regime dependent. The autoregressive order specifies the number of lags that are taken into account, while exogenous variable specifies the number and type of macroeconomic or PCA generated variables that are included.

Table 5.2: Model fit: Statistical performance

Model	Log-likelihood	BIC	Model	Log-likelihood	BIC
1	4628.995	-9220.156	12	4863.937	-9637.073
2*	4858.960	-9705.919	13	4861.384	-9647.099
3	4859.904	-9659.273	14	4860.473	-9645.278
4	4860.109	-9644.549	15	4770.781	-9496.161
5	4860.049	-9644.430	16	4774.806	-9489.078
6	4860.109	-9644.549	17	4786.094	-9481.386
7	4860.049	-9644.43	18	4788.793	-9486.783
8	4859.161	-9657.786	19	4804.253	-9517.704
9	4859.097	-9657.659	20	4789.150	-9487.498
10	4859.097	-9657.659	21	4798.803	-9476.538
11	4860.784	-9645.900			

*The table is the result of a fitting process on a dataset comprising of weekly returns on the MSCI World from 1987-01-01 to 2024-02-01. * indicates the best performing model specification in terms of BIC score.*

Parameter inference

It is remarkable that the model selection criteria scores of the different model specifications within the TCTP AR(1) class are almost identical. Whether the federal funds rate or the OECD CLI leading indicator is added to the model as an exogenous variable makes almost no difference. A potential explanation for this might be that the exogenous variables are not able to explain a lot about the model and are not significant. In order to test this hypothesis, the estimated parameters and their significance will be studied. The results of this analysis can be found in Table 5.3. Table 5.3 shows the regression coefficients for model specifications 3, 4, 9, and 17. The coefficient x1 refers to the first exogenous variable, which is equal to either the lagged value of the dependent variable y_{t-1} if included in the analysis or otherwise equal to the first exogenous variable. Furthermore, x2 is equal to the exogenous variable in case the model specification is AR(1). From looking at Table 5.3, it is clear that the volatility and the mean of the returns are significant variables, while the exogenous variables are not. This confirms the above suspicion that the exogenous variables hold no significance within the model. Furthermore, it should be noted that volatilities are the most significant variables, suggesting that these are the most relevant in determining the regime classification. This is in line with the literature (Guidolin and Timmermann, 2008).

Table 5.3: Model fit: Regression parameters

Coefficient	**Model**			
	3	**4**	**9**	**17**
Regime 1				
Intercept (μ)	0.0031***	0.0017***	0.0025***	0.0018***
x1	-0.0036	0.0012	-0.0023	0.0005
x2	-	-0.00052	-	0.0016
Volatility (σ)	0.0002***	0.0003***	0.0004***	0.0002***
Regime 2				
Intercept (μ)	-0.0028*	-0.0031*	-0.0026*	-0.0029*
x1	-0.0580	-0.0572	-0.0595	-0.0567
x2	-	0.00022	-	-0.0003
Volatility (σ)	0.0011***	0.0012***	0.0010***	0.0013***
Transition probabilities				
p[1->1]	0.9714***	0.9712***	0.9715***	-
p[2->1]	0.0651***	0.0648***	0.0653***	-

Notes: The table shows the regression coefficients for model specifications 3, 4, 9 and 17, for more details on the various models I refer to Table 5.1. The coefficient x1 refers to the first exogenous which is equal to either the lagged value of the dependent variable y_{t-1} if included in the analysis, or otherwise equal to the first exogenous variable. Lastly, x2 is equal to the exogenous variable in case the model specification is of AR(1).

Comparison to standard AR models

The results of the HMM analysis can be compared to those obtained from fitting standard linear models. In the process, the first research question of this paper will be answered:

1. Which model is the most efficient in capturing stylized effects commonly observed in the return distributions of an institutional investor portfolio over time?

In order to perform the aforementioned comparison, a set of linear models needs to be selected. This set of model specifications can be found in Table 5.4. The set of model specifications considered is mostly equal to the set of HMMs considered, since this enables the fairest comparison and is closest to the current models used at Sprenkels. It is clear from the comparison of Table 5.4 and Table 5.1 that the HMMs outperform the AR-model benchmark models. Implying that HMMs are better at capturing the time-varying effects and non-linear patterns displayed in asset returns. This is a crucial finding since it implies that standard, non-time-varying AR models might not be the optimal choice for modeling asset returns in the asset allocation models of Sprenkels.

Table 5.4: Overview of autoregressive models

Model	Exogenous variables	Log Likelihood	BIC
AR(0)	-	4628.995	-9242.857
AR(0)	OECD Leading indicator	4629.078	-9235.456
AR(0)	One Principal component	4629.155	-9235.609
AR(1)	-	4627.433	-9232.166
AR(1)	Term spread rate	4627.433	-9224.600
AR(1)	FED Funds rate	4627.466	-9224.667
AR(1)	OECD CLI Leading indicator	4627.517	-9224.769
AR(1)	One Principal component	4627.628	-9224.990
AR(1)	Two Principal components	4629.529	-9221.227

Notes: This table displays the he result of a fitting process on a dataset comprising of weekly returns on the MSCI World from 1987-01-01 to 2024-02-01.

Conclusion

Concluding, TCTP models score better on model selection criteria than their TVTP counterparts. Thus suggesting that they are better suited for inference. Furthermore, within the class of TCTP models, the differences between model specifications are minimal. These minimal differences within classes stem from the fact that the exogenous variables are not significant within the regression. Lastly, the HMM performs significantly better than the standard AR model, suggesting that there are non-linear patterns in the return series of assets that are effectively captured with the help of HMMs.

5.2 Forecasting Results

There are two distinct types of forecasting: in-sample forecasting and out-of-sample forecasting. In-sample forecasting involves training a model on a subset of observations and then forecasting the same subset of observations. On the flip side, out-of-sample forecasting involves training a model with one subset of the data (e.g., the training set) and attempting to forecast observations that are not part of this specific subset (e.g., the testing set).

For the specific task of forecasting stock market regimes and returns addressed in this paper, out-of-sample forecasting naturally holds greater relevance. Out-of-sample forecasting, commonly referred to as backtesting in the investment world, involves evaluating the performance of an investment strategy retrospectively (e.g., ex-post analysis). For this reason, this paper will focus on out-of-sample testing and only report the out-of-sample results. While backtesting is a crucial step in assessing the potential performance of an investment strategy, it is essential to recognize that past performance is no guarantee of future success. The essence of this principle is captured in the widely recognized stock market adage, "Past performance is not indicative of future results."

5.2.1 Out-of-sample regimes results

As previously mentioned in Chapter 2, there are two quantities that need to be forecasted. Firstly, the prevailing stock market regime, and secondly, the stock market returns. It makes sense to first discuss regime forecasts since these will be used in the return forecasts. The performance of the regime forecasts will be tested on their statistical accuracy as well as their economic value. The expectation is that these two metrics are correlated with each other and that higher statistical accuracy would lead to more economic value.

Statistical accuracy

Regime forecasting involves correctly identifying the occurrences of regime switches and, analogously, their persistence. First, the approach to obtaining the smoothed marginal probabilities will be discussed. These smoothed marginal probabilities represent the forecasts of future regimes. Please note that the results and outputs presented in this section are derived from the data specified in Chapter 3 and the methodology outlined in Chapter 4. Additional methodology will be introduced to properly evaluate the performance of the forecasts. The following is a brief overview of the aforementioned additional methodology. Firstly, a HMM with the chosen specification is selected. Once the model specification is chosen, it is used to forecast the data. The procedure employed for this involves a recursive forecasting approach with an expanding window.

The recursive forecasting procedure follows these steps: the entire dataset is divided into a training sample and a testing sample. The model is then fitted with the training sample. Subsequently, a recursive process for forecasting is initiated. At each step, the probabilities of being in regime 1 or 2 at time $t + 1$ are predicted. The key feature of the expanding window approach is that, with every step, the training set grows. Data points that have already been predicted are added to the training sample. This expansion means that, for the forecast at time-point T, the entire dataset up until and including $T-1$ is included in the training sample. Therefore, the forecasting procedure involves a sequential combination of forecasts, with each step contributing to the growing training set.

In order to assess the accuracy of our model's predictions, data regarding the various stock market regimes is required. This is problematic because the regime variable is an unobserved variable. This stems from the fact that there is no uniform definition of what exactly characterizes a bull or a bear market (Gonzalez et al., 2006). As stated in Chapter 4, the best possible option is to infer the actual state of the stock market by looking at the price development of the underlying assets. Despite the absence of a definitive consensus on the specific characteristics defining a bear or a bull market in stock prices, to facilitate the evaluation of forecasts, the decision is made to adopt the popular dating rule first introduced by Lunde and Timmermann (2004) and later used in studies from Kole and Van Dijk (2017); Haase and Neuenkirch (2023). The main advantage of this dating rule is that it does not make any assumptions about the underlying distribution of the return series. The structure of the dating rule is as follows:

Bull Market Classification

- Given that the last observed extreme was a local maximum, referred to as $P_{\max}$, the subsequent price series is checked against the following criteria:

  (a) $P_{\max}$ is updated if the stock market has risen above the last $P_{\max}$, and the time period is classified as a bull market.

  (b) A local minimum $P_{\min}$ has been found if the stock market has fallen by 15% or more with respect to the last $P_{\max}$, and the time series is classified as a bear market.

  (c) There are no updates if neither (a) nor (b) took place, and the time series is classified as a bull market.

Bear Market Classification

- Given that the last observed extreme was a local minimum, referred to as $P_{\min}$, the subsequent price serie is checked against the following criteria:

  (a) $P_{\min}$ is updated if the stock market has dropped below the last $P_{\min}$, and the time point in the series is classified as a bear market.

  (b) A peak has been found if the stock market has risen by 10% or more with respect to the last $P_{\min}$, and the time point in the series is classified as a bear market.

  (c) There are no updates if neither (a) nor (b) took place, and the time point is a bear market.

The regime division resulting from this dating rule is illustrated in Figure 5.1. It is evident from the figure that this dating rule effectively identifies major bear markets of the past decades, including the 2008 financial crisis, the 2000 Dotcom bubble, and the COVID-19 crisis in 2022. Simultaneously, it accurately identifies bull market periods.

Figure 5.1: MSCI World Price Series with Bear Market Periods Highlighted

Note: The Figure shows the MSCI World price index and the identified bear market periods as gray-shaded areas. The classification follows the dating rule of (Lunde and Timmermann, 2004).

The resulting division obtained from Lunde and Timmermann (2004) dating rule will serve as a benchmark for evaluating the performance of the regime forecasts. In order to determine the statistical performance of the various models, a set of performance parameters needs to be selected. In line with Haase and Neuenkirch (2023), the following variables have been chosen to evaluate the performance:

- **Accuracy**: the accuracy of our predictions is defined as the share of correct predictions in the total number of predictions. However, in order to evaluate this metric, a threshold value with which probabilities can be converted to actual binary predictions is needed. For now, the threshold value is set at 0.5, which at first glance seems most intuitive and is in line with the literature (Haase and Neuenkirch, 2023). The total accuracy as well as the bear market and bull market classification accuracies will be reported.

- **Area under curve (AUC)**: the AUC is a commonly employed metric in the context of evaluating the performance of binary classification models. It represents the entire area under the receiver operating characteristic (ROC) curve. The ROC curve plots the true positive rate against the false positive rate for each possible threshold value. The threshold value dictates whether or not a probability is converted to a bear market or a bull market. For accuracy, the threshold value is set equal to 0.5 because, ex ante, it is not known what the optimal threshold value is. AUC studies the entire range from 0 to 1. For the AUC, a higher value indicates better discrimination between bull and bear regimes. An AUC score of 0.5 basically means that the model does not perform better than random guessing, while an AUC of 1 represents perfect classification. Figure 5.2 displays an example of such a ROC curve to give some more intuition behind the concept.

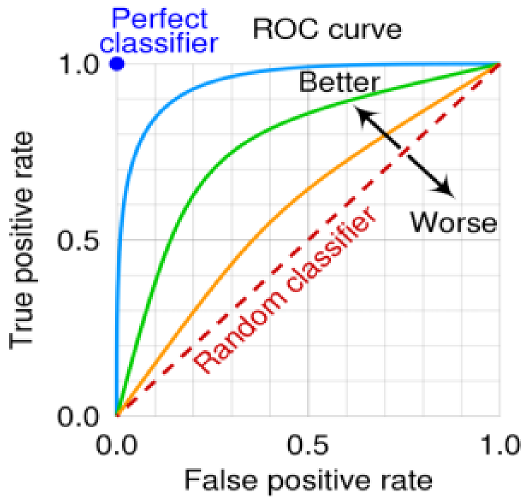

Figure 5.2: Example ROC curve (Campion and Campion, 2023)

- **Quadratic probability score (QPS)**: the QPS is a metric that works well for the evaluation of probabilistic forecasts, particularly in cases of binary classification. It measures the difference between the forecasted probabilities and the real values. The QPS punishes both overconfident and underconfident forecasts. The QPS is calculated as the sum of the squared differences between the predicted probabilities and the real values for each observation, then divided by the total number of observations. A lower QPS value indicates better forecast performance.

The results of the forecasting analysis for the time period from 2012-04-01 to 2024-02-01 can be found in Table 5.5. From Table 5.5, a few interesting insights can be derived regarding the statistical accuracy of the predictions made by the model specifications. The best-performing model in terms of both the QPS and the AUC score is model specification 11. This model belongs to the TCTP AR(1) class and includes one PCA-generated factor as an exogenous variable. Similarly to the results obtained in Section 5.1, it can be observed that there are no significant differences between the different AR(1) models with exogenous variables in terms of statistical performance within the TCTP class.

The introduction of TVTP models does not improve the statistical performance of the forecasts, as is evident from a significant decrease in the AUC score. This drop in the AUC score coincides with an increase in the accuracy of identifying bear market regimes, indicating that the model classifies a larger number of periods as bear markets at the expense of misclassifying bull market periods as bear markets, also known as false positives. Furthermore, basic TCTP models like model specifications 1 and 2 do not perform well in terms of statistical accuracy, contradicting the earlier findings in Section 5.1, where model specification 2 ultimately was deemed the best performing model. This result seems counter-intuitive since it is generally expected that simpler models will perform better in terms of forecasting accuracy but worse in terms of model fit, and vice versa. However, the correlation between model complexity and forecasting capability is not linear, and it is entirely plausible that greater model complexity could result in enhanced forecasts.

Moreover, it should be noted that the above findings partly align with conclusions drawn in Section

5.1, where it was already noted that there were no significant differences within particular classes of models. This suggests that the performance of the models may not rely strongly on external factors. There are slight differences between the findings made in this paper and earlier findings by Haase and Neuenkirch (2023), since they have shown that there are TVTP models that are able to outperform standard TCTP models. However, this can be attributed to differences in the exact model specification selection and differences in the research set-up.

Table 5.5: Regime forecasts: Statistical performance

Model number	QPS	AUC	Accuracy	Bear	Bull
1	0.2501	0.4923	0.3776	0.5294	0.3688
2	0.2577	0.6678	0.6888	0.5490	0.7014
3	0.0949	0.8178	0.8752	0.4706	0.9117
4	0.0936	0.8197	0.8768	0.4706	0.9134
5	0.0917	0.8257	0.8817	0.4706	0.9187
6	0.0932	0.8200	0.8801	0.4314	0.9205
7	0.0917	0.8257	0.8817	0.4706	0.9187
8	0.0932	0.8200	0.8801	0.4314	0.9205
9	0.0932	0.8200	0.8801	0.4314	0.9205
10	0.0949	0.8178	0.8752	0.4706	0.9117
11*	0.0914	0.8277	0.8801	0.4706	0.9170
12	0.0921	0.8247	0.8768	0.4314	0.9170
13	0.0944	0.8192	0.8768	0.4706	0.9134
14	0.0943	0.8195	0.8768	0.4510	0.9152
15	0.2359	0.6492	0.6759	0.5490	0.6873
16	0.2352	0.6465	0.6759	0.5490	0.6873
17	0.2156	0.6701	0.7245	0.5294	0.7420
18	0.2171	0.7176	0.6904	0.5882	0.6996
19	0.2599	0.6965	0.6515	0.6471	0.6519
20	0.2687	0.6791	0.6499	0.6078	0.6537
21	0.2128	0.7122	0.7196	0.5686	0.7332

Notes: Models 1-14 belong to the TCTP class of models, while models 15-21 belong to the TVTP class of models. The QPS is defined as the quadratic probability score. AUC is defined as the area under the curve. The accuracy is defined as the share of correctly predicted regimes, where the smoothed marginal probabilities are converted into binary values through a threshold value of 0.5*. Bear and Bull are the share of correctly predicted bearish and bullish regimes respectively.* indicates the best performing forecast in terms of the AUC*

As indicated earlier, model Specification 11 is the best-performing model based on its QPS and AUC score. Therefore, a more detailed analysis of its results will be conducted. The top graph in Figure 5.3 illustrates the smoothed marginal probabilities of being in either a bull or bear stock market regime for model Specification 11. In contrast, the bottom graph of the same figure displays the benchmark values obtained from the dating rule for being in a bull or bear stock market regime. From an initial examination, it appears that the model can effectively anticipate significant bear market periods, although with a slight delay. From the top graph of Figure 5.3, it can also be deduced that a threshold value of 0.5 might not be optimal for the model. Stemming from the fact that there are several instances in which the model transitions to a bear market while the actual model does not warrant a transition. However, incorporating that would make the analysis in-sample.

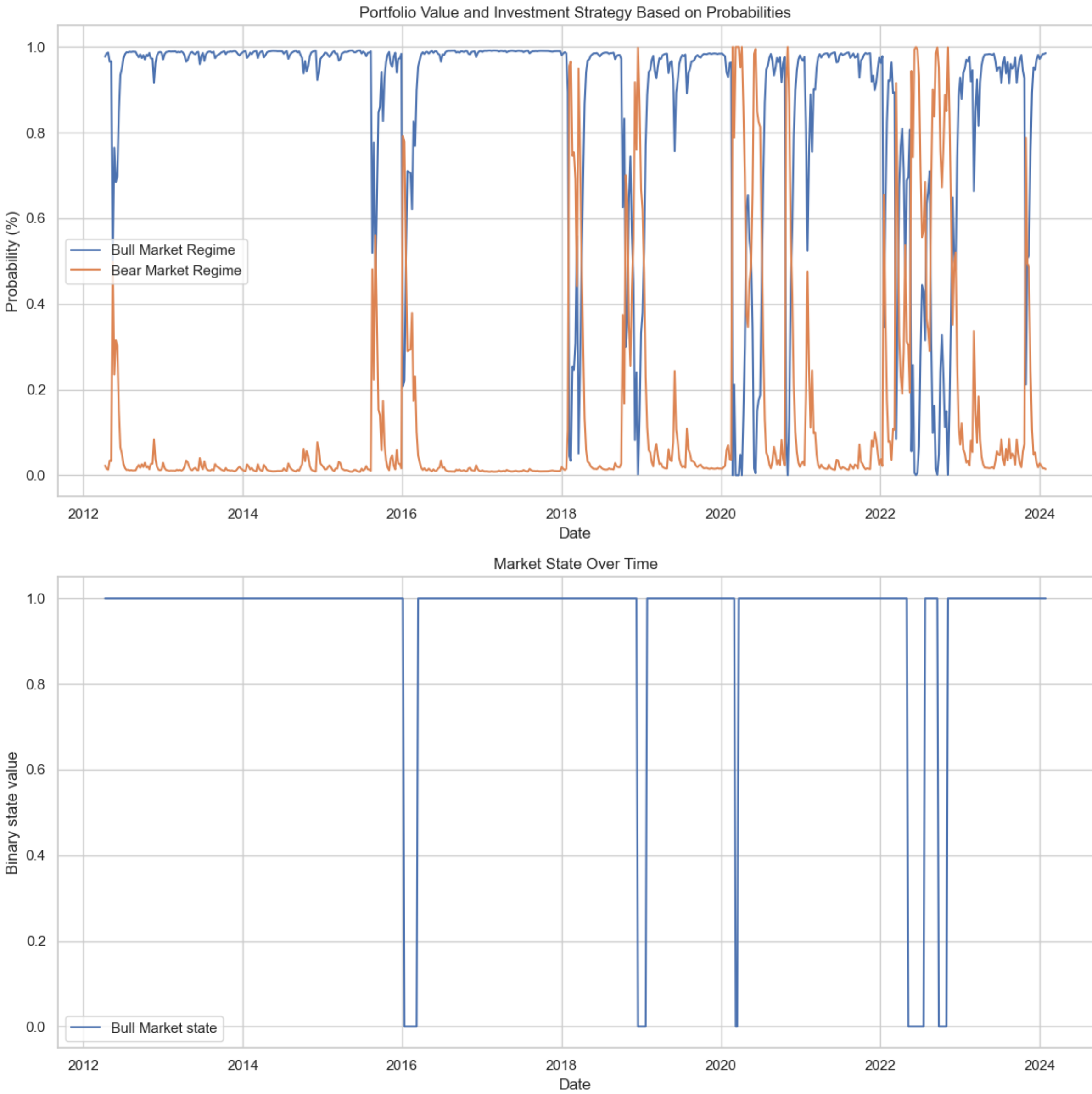

Figure 5.3: Smoothed marginal probabilities and the actual regime classification

Note: The figure reports the smoothed marginal probabilities of model specification 11 for the time period 2012-04-01 to 2024-02-01.

While it is interesting to look at the classification ability of the models in terms of bull or bear markets, it is also essential to discuss the models' ability to predict regime switches. These regime switches serve as a signal of a possible market shift. Missing shifts might be costly in terms of losing possible hedging opportunities. While signaling a regime switch too often can invalidate the models' use due to a possible erosion of trust and possibly trading costs incurred in the process of preparing for such a switch. Table 5.6 presents a summary of all regime switches observed throughout the sample period, along with information regarding the model's ability to forecast these switches. The model is considered successful in predicting a switch if it also transitions within a month before or after the actual switch. From Table 5.6, it can be inferred that the HMM performs relatively well in predicting switches from bull market regimes to bear market regimes. However, its performance is notably weaker in predicting switches from bear market regimes to bull market regimes, which is in line with the literature (Ang and Timmermann, 2012; Haase and Neuenkirch, 2023). This discrepancy is likely associated with the HMMs' effectiveness in forecasting structural breaks, which aligns more closely with switches from bear to bull markets, seeing as these are historically more abrupt.

Table 5.6: Regime switching prediction results

Date	Prediction	Type of switch	Prediction delay
2016-01-11	True	Bull to Bear	-1
2016-03-14	False	Bear to Bull	-
2018-12-17	True	Bull to Bear	-3
2019-01-28	True	Bear to Bull	-1
2020-03-09	True	Bull to Bear	-2
2020-03-23	False	Bear to Bull	-
2022-05-09	True	Bull to Bear	-2
2022-07-25	True	Bear to Bull	1
2022-09-26	False	Bull to Bear	-
2022-11-07	True	Bear to Bull	3

Notes: The table shows the regime prediction performance of model specification 11. A prediction is true in case the prediction delay is within a month of the actual switch. Furthermore, the prediction delay is reported in weeks. The dating rule of Lunde and Timmermann (2004) is used for the benchmark classification of bull and bear markets.

In order to further analyze the performance of our models, a confusion matrix is used. A confusion matrix is an effective tool to analyze all the important aspects of the model's forecasts since it takes into account not only the accuracy of a prediction but also its false positive rate. As noted, one could conclude that the model performs relatively well in predicting switches from a bull market regime to a bear market regime from looking at the results in Table 5.6. Nevertheless, Table 5.7 depicts a contrasting situation. Although 4 out of 5 actual switches are predicted correctly, a total of 18 switches are predicted. This means the model will frequently sound a false alarm. This could result in a firm making trades that are not required, which results in extra transaction costs. To answer the question of whether the extra costs incurred from trading too often are outweighed by the benefits incurred from incorporating regime switches, an analysis of the economic value of the model forecasts will be done.

Table 5.7: Confusion matrix of regime switching results

Regime changes		Actual		
		Bear to Bull	Bull to Bear	None
Predicted	Bear to Bull	3	0	15
	Bull to Bear	0	4	14
	None	2	1	578

Notes: The table reports the confusion matrix of the regime switch predictions of model specification 11, the model can either predict a switch from Bear to Bull, Bull to Bear or no switch at all. The time period considered is 2012-04-01 to 2024-02-01 and there are 617 observations in the sample.

Economic value

Although the statistical accuracy of the models' forecasts is important, their economic value might be even more valuable. In order to determine the economic value of the model forecasts, an investment strategy will be developed that incorporates the regime forecasts, and the performance of that strategy will be compared to some common benchmarks.

Up until now, much of the literature discussion surrounding the evaluation of economic value has focused on the performance of HMM-based investment strategies within a relatively small investment universe of only one or two assets such as the works of Haase and Neuenkirch (2023); Ang and Timmermann (2012). Institutional investors that make up Sprenkels' clients, however, invest in a wider array of assets. Trying to fit a HMM on a data frame containing the return series of the entire investment universe is a complicated process. Complexity arises from the high dimensionality associated with regime-switching models that are fitted to that many variables. The high dimensionality would require a lot of data to get significant parameter estimates. Moreover, these calculations are deemed too complex and elaborate for this study. In the work of Nystrup et al. (2018), an efficient workaround to this problem has been found. This workaround fits with the investment goals of the institutional investors that are clients of Sprenkels. The approach outlined in the work of Nystrup et al. (2018) is based on the regime forecasts produced by the HMM. The basic idea is that based on the regime forecasts of the HMM, one allocates part of his assets to a risk-on/risk-off investment strategy. More specifically, the investment strategy operates in the following way:

According to a set amount p, a share of the portfolio of an institutional investor is allocated to the risk-on/risk-off portfolio, while the other share of the portfolio is allocated to the "normal" strategic asset allocation that lays somewhere in between the risk-on and the risk-off allocation. For now, let's assume $p = 0.5$, meaning that 50% of the total portfolio is invested in the risk-on/risk-off portfolio while the other half is invested in the strategic long-term asset allocation. All of the investment strategies mentioned above invest in the same asset universe, and their weights can be found in Table 5.8. The share of the institutional investors wealth allocated to the risk-on/risk-off portfolio will be put towards either the risk-on or the risk-off portfolio based on the regime forecasts that were obtained from the procedure that was first outlined in Section 5.2.2. More specifically, a regime switch to a bear market classification will see the allocation shift to the risk-off portfolio, and vice versa.

The input return series used to determine the regimes is the MSCI World, similarly to the work of Nystrup et al. (2018). The costs incurred for switching between the two strategies are set at a considerable amount, namely 20 basis points of the total amount of capital under management. This decision was made in accordance with experts in the business. As stated, Nystrup et al. (2018) is the first and, to our knowledge, the only one to implement the aforementioned investment strategy. In their paper they focus on proving the potential gains that could be made with the implementation of such a strategy while acknowledging that there is still a lot of potential extra gain to be made by further researching which type of HMM would yield the best returns in such a framework. In this paper, a wide array of HMMs is tested, and thus a significant contribution is made to the literature.

Table 5.8: Asset allocation weights

Name	Category	SAA	Risk-on	Risk-off
MSCI World	Stocks	25%	33.3%	12.5%
MSCI EM	Stocks	5.0%	6.7%	2.5%
Vanguard Real Estate Index Fund	Real estate	10.0%	13.3%	5.0%
VanEck Gold Miners ETF	Commodity	5.0%	6.7%	2.5%
United States Natural Gas Fund, LP	Commodity	5.0%	6.7%	2.5%
iShares iBoxx $ High Yield Corporate Bond ETF	Credit	5.0%	6.7%	2.5%
iShares J.P. Morgan EM High Yield Bond ETF	Credit	5.0%	6.7%	2.5%
	Subtotal	60%	80%	30%
iShares 20+ Year Treasury Bond ETF	Fixed income	20%	10%	35%
iShares 1-3 Year Treasury Bond ETF	Fixed income	20%	10%	35%
	Subtotal	40%	20%	70%

Note: These are the weights assigned to the different assets for the SAA portfolio, the risk-on portfolio and the risk-off portfolio.

With the investment strategy laid out, the next step is to define the performance metrics that will be used to judge the performance of the various model specifications and their benchmarks. For the benchmark, four different investment strategies have been chosen, where the investor either invests 100% of his wealth in the strategic asset allocation, the risk-on allocation, the risk-off allocation, or the MSCI World. When it comes to assessing the performance of an investment strategy, one does not only look at the performance in terms of absolute returns but also at the volatility of the strategy:

- $\mathbf{R^{CUM}}$ = represents the cumulative wealth generated by the investment strategy, assuming an initial investment of $1.
- $\hat{\mathbf{R}}$ represents the average annual return of the investment strategy.
- $\hat{\boldsymbol{\sigma}}$ = measures the annualized volatility or risk of the investment strategy.
- **Sharpe Ratio (SR)** = measures the risk-adjusted return. Calculated as the ratio of the average excess return to the standard deviation of returns.
- **Certainty Equivalent Return (CER)** = measures the return that an investor with a certain level of risk aversion (γ) would find equally desirable as the risky return. $CER = \hat{R} - \frac{\hat{\sigma}^2}{2} \cdot \gamma$, where the value of γ is set to 3, in line with the literature (Haase and Neuenkirch, 2023).
- **Maximum drawdown (MaxDD)** = represents the maximum loss from a peak to a trough of an investment strategy before a new peak is attained. It is a measure of the largest observed loss over a specified period.

Now that the investment strategy and the performance metrics are defined, it is possible to start analyzing the economic value of our models. The entire set of models that is defined in Table 5.1 will once again be tested, since this will allow us to analyze the relationship between model fit, statistical accuracy, and the model's ability to generate economic value. The time period considered in this forecasting analysis is from (2012-04-01 to 2024-02-01), and the results can be found in Table 5.9.

Table 5.9: Regime forecasts: Economic value.

Benchmark	R^{cum}	$\hat{R}$	$\hat{\sigma}$	SR	CER	MaxDD
SAA	1.527	0.009	0.043	0.062	0.008	-0.428
Risk-On	1.631	0.011	0.054	0.059	0.010	-0.475
Risk-Off	1.352	0.006	0.032	0.055	0.006	-0.355
MSCI World	2.544	0.021	0.074	0.082	0.008	-0.643
Model (TCTP)						
1	0.124	-0.039	0.044	-0.258	-0.040	-0.887
2	1.597	0.010	0.035	0.081	0.009	-0.483
3*	1.791	0.012	0.041	0.086	0.011	-0.544
4*	1.813	0.013	0.041	0.088	0.012	-0.543
5*	1.794	0.012	0.041	0.086	0.011	-0.543
6*	1.813	0.013	0.041	0.088	0.012	-0.543
7*	1.794	0.012	0.041	0.086	0.011	-0.543
8*	1.819	0.013	0.042	0.087	0.012	-0.557
9*	1.819	0.013	0.042	0.087	0.012	-0.557
10*	1.791	0.012	0.041	0.086	0.011	-0.544
11*	1.803	0.012	0.041	0.087	0.012	-0.543
12*	1.875	0.013	0.041	0.092	0.012	-0.558
13*	1.756	0.012	0.041	0.083	0.011	-0.536
14*	1.804	0.012	0.042	0.086	0.012	-0.550
Model (TVTP)						
15	1.341	0.006	0.035	0.051	0.006	-0.398
16	1.404	0.007	0.035	0.058	0.006	-0.412
17	1.308	0.006	0.036	0.046	0.005	-0.376
18	1.426	0.007	0.037	0.058	0.007	-0.459
19	1.437	0.008	0.035	0.061	0.007	-0.456
20	1.494	0.008	0.035	0.068	0.008	-0.458
21	1.382	0.007	0.038	0.052	0.006	-0.411

*Notes: The table displays the economic value of various investment strategies. SAA, risk-on and risk-off correspond to the weights defined in Table 5.8. MSCI World corresponds to an 100% investment in the MSCI World. Models 1-14 belong to the TCTP class of models, while models 15-21 belong to the TVTP class of models. R^{CUM} is the cumulative wealth of the strategy, with a starting investment of €1, $\hat{R}$ are the annualized average returns, $\hat{\sigma}$ are the annualized standard deviations, SR is the Sharpe ratio, CER is defined as the certainty equivalent return with γ set to 3. MaxDD is the maximum drawdown. More information on the different criteria can be found in Section 5.2.1. * here, indicates models that outperform the benchmarks in terms of risk-adjusted returns as measured by the SR and the CER.*

From the figures in Table 5.9, a few interesting conclusions can be drawn. In terms of absolute returns, no strategy outperforms the 100% MSCI World benchmark strategy. However, such a strategy would not fit the risk appetite of the institutional investors that Sprenkels typically consults, due to its high volatility. When the MSCI World strategy is excluded from the analysis, the models belonging to the TCTP AR(1) class once again perform best in terms of risk-adjusted returns. Furthermore, it is clear that those models significantly outperform most of the benchmark strategies on risk-adjusted and absolute returns. While additionally outperforming the MSCI World on a risk-adjusted basis.

In terms of absolute returns and risk-adjusted returns, model specification 12 performs best in this specific sample, meaning that a strategy implementing the forecasts of an TCTP AR(1) model with two exogenous PCA-generated factors leads to the best economic value. This is not surprising since this model also performs well when it comes to its statistical accuracy in predicting regimes. In order to check whether this assessment holds, a scatter plot was created that graphs the model's statistical accuracy relative to its economic performance. The result of this can be found in Figure 5.4.

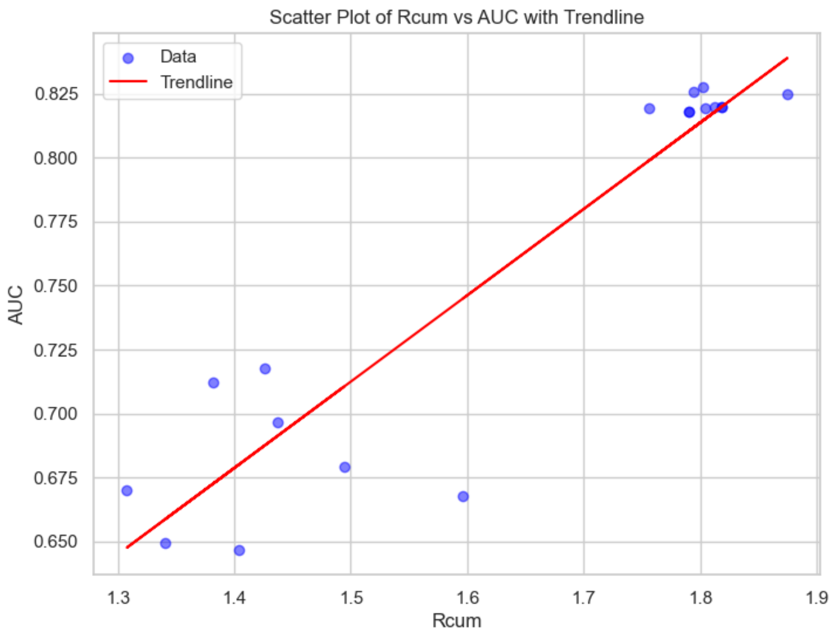

Figure 5.4: Scatterplot of the economic value measured in cumulative returns and the statistical accuracy as measured by the AUC score.

As expected, worse statistical performance seems to imply worse economic performance, as can be seen in Figure 5.4. However, the integration of the costs related to switching between the risk-on and the risk-off portfolios is what largely determines this relationship. The costs of switching between the risk-on and the risk-off portfolios are currently 0.2% of the total portfolio every time a switch is made.

Next to the analysis of various performance metrics, it is also interesting to study the cumulative wealth of an investor's portfolio. Figure 5.5 shows the cumulative wealth of the various benchmarks and that of strategy based on model specification 8. It is clear from Figure 5.5 that the HMM-based strategy outperforms almost all of the benchmarks significantly in terms of cumulative wealth. This is mainly due to the increased allocation of stocks during bull market periods.

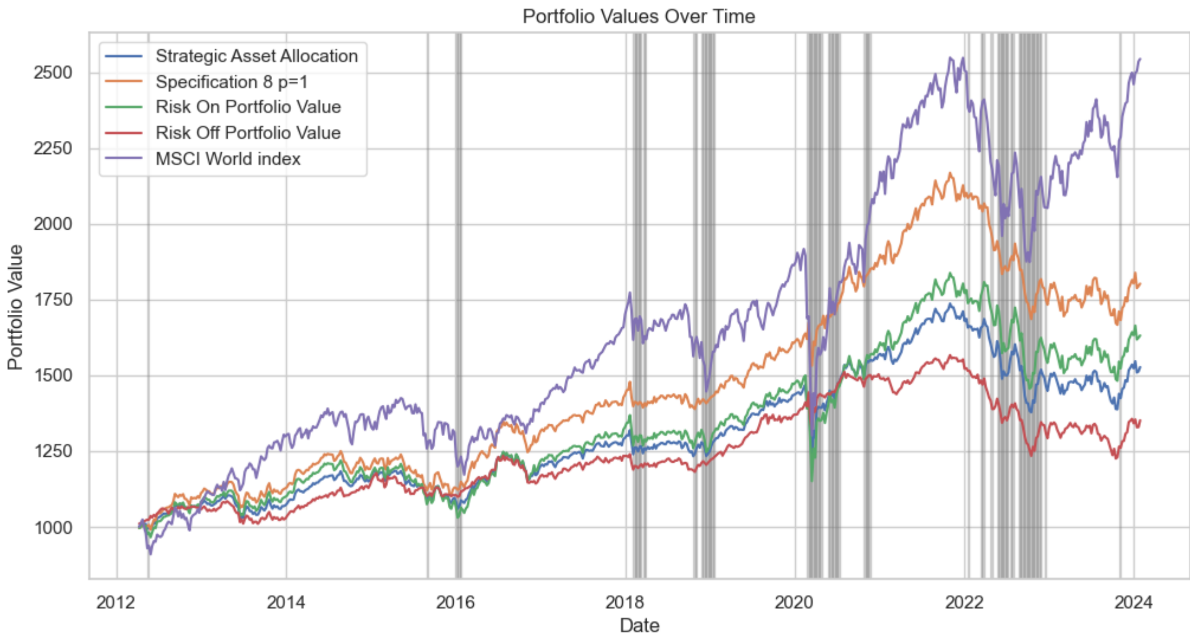

Figure 5.5: Development of cumulative economic value for model specification 8 and the selected benchmarks for the time period from 2012-04-01 to 2024-02-01.

Notes: The starting wealth for all investment strategies is €1000, and the portion of wealth allocated p to the risk-on/risk-off portfolio is 1*. Furthermore, the gray-shaded areas correspond to periods that are classified as a bear market by model specification 8.*

When considering the economic value of an investment strategy, it is important to also consider the practical consequences of actually implementing that strategy. This paper will provide a concise overview of the key aspects and how they are addressed. The normal transaction costs that might be incurred in order to make sure that the portfolio weights are respected are not taken into account. Partly because these costs are incurred with every strategy, which means in terms of our analysis, there is no real difference. Mapping errors are assumed to be negligible, which is reasonable since the effects of such errors are similar for all of the strategies studied, including the benchmarks, so they will not affect the results of the analysis in a meaningful way. Furthermore, all of this is under the assumption that the threshold is equal to 0.5. So, the best possible return that is attainable might be different.

One of the more problematic implications has to do with the implementation of a delay in portfolio switches. Due to the structure and size of the average institutional investor, decisions on the relocation of funds typically require several weeks of throughput time. This is problematic, since taking into account such delays means that gains in returns all but disappear. The results behind this analysis can be found in Appendix 9.4. However, the results in Table 5.9 suggest that these strategies are an attractive option that should be investigated further by Sprenkels since structural changes to the governance structure that would make implementation of HMM-based strategies possible could lead to significant capital gains.

Conclusion

The analysis of the various HMM specifications in terms of their statistical accuracy and their economic value yields several key findings. Firstly, the model specifications belonging to the TCTP AR(1) class seem to perform best on both aspects, outperforming their TVTP counterparts. Furthermore, there exists a relationship between the statistical accuracy of the models and the economic value they generate, with models exhibiting better statistical performance generally generating superior economic outcomes. When looking at the statistical accuracy of even the best-performing model specifications, the conclusion can be drawn that there is still a lot of improvement needed since even the best specifications tend to predict far too many switches.

This also ties into the significance of the transaction costs that are associated with switching between risk-on and risk-off portfolios on the economic value generated by the various strategies. Furthermore, there are several practical implications that make the implementation of HMM-based strategies difficult, like the throughput time associated with switching an investment strategy. It is promising, however, that when compared to traditional benchmarks, such as the strategic asset allocation and a market index like the MSCI World, the regime-based strategy shows outperformance in terms of risk-adjusted returns. Further analysis, including sensitivity testing of threshold values like in Nystrup et al. (2018), could provide additional insights into optimizing the strategy for real-world implementation.

5.2.2 Out-of-sample return results

After having discussed the results of the out-of-sample regime forecasts, these results can now be used to predict returns on the MSCI World index for the out-of-sample period. These MSCI World return predictions are then utilized to derive weights for mean-variance optimization, enabling us to make asset allocation decisions. Assessing the effectiveness of the return forecasts involves evaluating their economic value and their statistical accuracy. Much of the same process that was first laid out in Section 5.2.1 will be followed in this section.

This analysis serves two main purposes. Firstly, it allows us to evaluate the statistical performance of the selected HMMs in terms of their ability to predict returns. Predicting returns is particularly interesting because, unlike regime forecasting, the actual quantity being forecasted is known and observable. Therefore, it is possible to objectively assess the quality of the forecasts against the actual realized returns. Additionally, the goal of this section is to ascertain whether strategies integrating return forecasts can yield substantial economic outperformance compared to a certain set of benchmark strategies.

Moreover, by conducting a statistical and economic value analysis of the return forecasts, it becomes possible to compare these results with the regime forecasting results obtained in Section 5.2.1. Evaluating whether accurate regime forecasts translate into reliable return forecasts provides valuable insights into the quality of the selected regime classification benchmark and the suitability of the testing procedure outlined earlier in Section 5.2.1. This additional step is also a new contribution to the existing literature, since cross-checking the results of the return forecasts with those of the regime forecasts has not been done yet to our knowledge.

Statistical Accuracy

In order to discuss the statistical accuracy of the return forecasts of the various model specifications, two key elements are needed: a procedure that enables us to obtain return forecasts and performance metrics to judge the quality of the forecasts. First, the former will be discussed. The forecasting procedure employed in this section builds on the procedure that was first described in Section 5.2.1. As stated in Section 5.2.1, a recursive forecasting procedure is used, where at each time step t, the probabilities of being in a bear market regime or a bull market regime at time $t+1$ are predicted. Then, by using equation (4.16), the future values of the MSCI World returns are predicted given the value of the state variable s_t. These predicted values are then weighed with the probabilities of being in a bear market regime or a bull market regime, which yields return forecasts on the MSCI World, independent of the state variable s_t. This procedure is similar to the one employed by Haase and Neuenkirch (2023). In order to evaluate the performance of the forecasts, another set of performance metrics is needed. Because the previously stated set of performance metrics in Section 5.2.1 are not useful in this context since the type of forecasting being done in this section is not of the binary classification type. The following three performance metrics are used in the analysis:

- $$MSFE = \frac{\sum(\hat{y}_t - y_t)^2}{n}. \tag{5.2}$$

  The mean squared forecast error (MSFE) measures the average squared difference between the forecasted values and the actual values. It provides an assessment of the forecasting accuracy by considering both the magnitude and direction of errors. Please note that a lower MSFE indicates better performance in terms of forecasting.

- $$RSME = \sqrt{\frac{1}{n}\sum(\hat{y}_t - y_t)^2}. \tag{5.3}$$

  The root mean squared error (RMSE) is the square root of the MSFE; it provides a measure of the typical deviation of the forecast errors from the actual values. It is easily interpretable and provides insight into the average magnitude of the forecasting errors. Similar to MSFE, a lower RMSE means better forecasting accuracy.

- $$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}. \tag{5.4}$$

  The Coefficient of Determination (R^2) measures how well the statistical model is able to predict returns. It indicates how well the predicted values explain the variability in the actual returns. A higher R^2 suggests a better fit of the predicted values to the actual data.

Table 5.2 displays the statistical performance of the return forecasts of all of the model specifications defined in Table 5.1 for the time period from 2012-04-01 to 2024-02-01. From the results presented in Table 5.10, it becomes clear that model specification 2, corresponding to the **TCTP**, **AR**(1) **with Switching Variance**, performs the best among all model specifications. Model specification 2 is one of the simplest models, so its above-average performance in forecasting was expected since simpler models tend to perform better in terms of forecasting. However, the overall level of the MSFE and the RMSE is relatively high, indicating poor return forecasting performance. A RMSE of around 2% is notably high when one compares it to actual weekly returns, which typically range between -2% and +2%. Furthermore, the R^2 results are not sufficient, falling significantly below the acceptable threshold of 0.5. These findings suggest that the statistical performance of the return forecasts is unsatisfactory.

There are two plausible explanations for this conclusion. Firstly, it is possible that the regime forecasts themselves are inaccurate and that these inaccurate forecasts lead to poor return forecasts. Alternatively, translating regime forecasts into return forecasts involves an additional challenging step. Point-forecasting of specific values is often more difficult than forecasting regimes. Hence, it is plausible that either the regime classification benchmark is flawed or that insignificant results are produced because of the return forecast step. Further investigation is needed to ascertain the underlying cause of these insignificant results.

Table 5.10: Return forecasts: Statistical performance

Model	MSFE	RMSE	R^2	Model	MSFE	RMSE	R^2
1	0.00046	0.02136	0.00054	12	0.00045	0.02126	0.00992
2*	0.00045	0.02111	0.02363	13	0.00046	0.02141	-0.00451
3	0.00046	0.02132	0.00424	14	0.00046	0.02134	0.00214
4	0.00045	0.02122	0.01395	15	0.00045	0.02116	0.01934
5	0.00045	0.02125	0.01070	16	0.00045	0.02122	0.01313
6	0.00045	0.02125	0.01070	17	0.00045	0.02123	0.01300
7	0.00045	0.02125	0.01070	18	0.00045	0.02133	0.00352
8	0.00045	0.02119	0.01642	19	0.00045	0.02126	0.01014
9	0.00045	0.02119	0.01642	20	0.00045	0.02123	0.01274
10	0.00046	0.02132	0.00424	21	0.00045	0.02123	0.01215
11	0.00045	0.02124	0.01119				

*Notes: The table displays the statistical accuracy as measured by the MSFE, the RMSE and the R^2. Models 1-14 belong to the TCTP class of models, while models 15-21 belong to the TVTP class of models. * here, indicates the best performing model in terms of R^2.*

In order to determine the root cause for the poor statistical performance of HMM-based models in terms of predicting returns, I will analyze a graphical overview of the return forecasts. Figure 5.6 displays the return forecasts of model specification 2 as well as the real data for the entire sample. From Figure 5.6, it is evident that the forecasts fail to accurately capture the magnitude of the real data. Even when bear markets are accurately forecasted, the model still fails to accurately match the magnitude of the data. Suggesting that the root cause does not lie with the regime predictions. Furthermore, from the results in Table 5.10 and Figure 5.6, it is obvious that the models are not able to effectively forecast returns on the MSCI World, and the quality of the forecasts relative to the performance of the regime forecasts is worse. An interesting question, however, is how the various model specifications perform in terms of economic value.

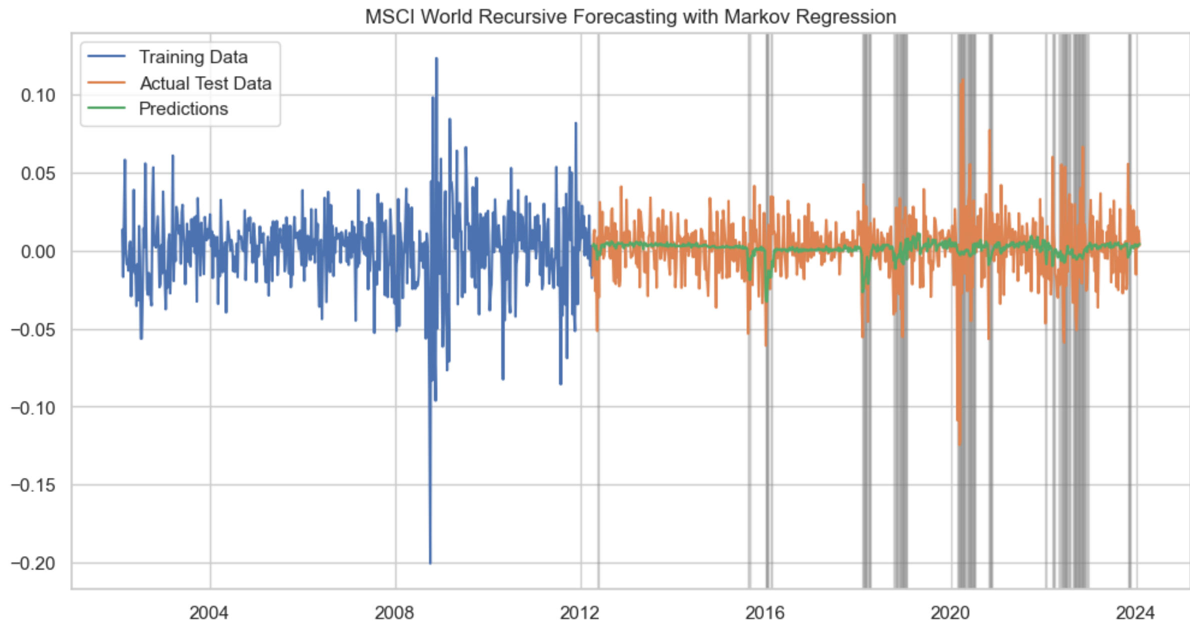

Figure 5.6: Asset return forecasts on the MSCI World produced by model specification 2 and the real returns on the MSCI World for the time period from 2012-04-01 to 2024-02-01.

Note: The gray-shaded areas correspond to periods that are classified as a bear market by model specification 2.

Economic value

In order to evaluate the economic value of the return forecasts, three elements are required: a trading strategy translating the return forecasts into economic value, benchmarks on which the performance of the alternative strategies is compared, and performance metrics upon which the comparison can be quantified. For the performance metrics, the choice was made to use the same set of metrics that was used in Section 5.2.1. In this paper, the following benchmark portfolios will be considered that are commonly used in the literature:

- **Buy and Hold Portfolio** A portfolio comprising of 60% marketable securities (represented by the MSCI World Index) and 40% fixed-income securities (represented by the iShares 20+ Treasury Bond ETF). The buy-and-hold portfolio strategy is a straightforward portfolio strategy where the initial asset allocation is maintained throughout the investment horizon.
- **100% Investment in the MSCI World:** A widely recognized benchmark, the MSCI World represents global equity performance.
- **Mean-variance Optimization (MVO) Portfolio:** This portfolio allows the investor to allocate funds to two asset categories, namely marketable securities (represented by the MSCI World Index) and fixed-income securities (represented by the iShares 20+ Treasury Bond ETF). The investor rebalances the portfolio on a weekly basis based on mean-variance optimization using the past 10 years of data. The return estimates for mean-variance optimization are obtained from a AR(1) model without regime switching.

With the benchmark portfolios and the performance metrics defined, the last element that is needed for the evaluation of the economic value is a trading strategy that translates the return forecasts of the HMMs into economic value. The trading strategy considered in this paper is based on mean-variance optimization and will loosely follow the work of Haase and Neuenkirch (2023). For mean-variance optimization to work, forecasts of mean returns and volatilities of the considered assets are required. In Section 5.2.2, the recursive procedure responsible for producing mean return forecasts can be found. The volatility of the MSCI World is predicted by calculating estimates using historical averages of the volatility in different regimes, which are then adjusted depending on the probability of being in those regimes. Formally:

$$
\begin{aligned}
E\left(\sigma_{t+1} \mid \Omega_t; \boldsymbol{\theta}\right) &= \sum_{j=1}^{2} \Pr\left\{s_{t+1} = j \mid \Omega_t; \boldsymbol{\theta}\right\} \cdot E\left(\sigma_{t+1} \mid s_{t+1} = j, \Omega_t; \boldsymbol{\theta}\right) \\
&= \sum_{j=1}^{2} \hat{\xi}_{t+1|t} \cdot \sigma_{s_t}
\end{aligned}
\tag{5.5}
$$

Based on these volatility and return forecasts, the optimal weights that should be assigned to marketable securities according to mean-variance optimization can be found:

$$
W_t^* = \frac{1}{\gamma} \frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2} \tag{5.6}
$$

The share of his wealth an investor can invest in marketable securities is capped at 1 and should be at least 0, which are reasonable assumptions considering institutional investors like pension funds typically avoid short positions and taking on excessive leverage. The remaining funds, i.e., the difference between the weight of the portfolio invested in marketable securities and the investor's entire wealth, are invested in fixed-income securities. Multiplying the weights for the two asset categories with the returns on both yields the portfolio return. Transaction costs are included by incurring a cost of 20 basis points of the entire wealth every time a trade is made that shifts the weights by more than 0.5 and 10 basis points every time a trade is made that shifts the weights less than 0.5 but more than 0.1 to rebalance the portfolio.

The economic value of the HMM-based strategies and benchmarks can be found in Table 5.11, the time period to be forecast in this analysis is from 2012-04-01 to 2024-02-01. The strategies all assume no extra investments or withdrawals. The weights suggested might lead to fractional shares, which may not be entirely realistic for individual investors. However, this assumption is considered appropriate, given the fact that the wealth of an institutional investor is at least a \$100 million. As expected, of the benchmark strategies, the 100% MSCI World investment strategy generates the largest absolute return but performs worse in terms of risk-adjusted returns when compared to the mean variance optimization benchmark strategy. The 60/40 benchmark strategy performs worst in terms of (risk-adjusted) returns.

Table 5.11: Return forecasts: Economic value

Benchmark	R^{cum}	$\hat{R}$	$\hat{\sigma}$	SR	CER	MaxDD
MVO Port.	2.303	0.022	0.070	0.065	0.012	-0.645
100% MSCI World	2.503	0.021	0.074	0.060	0.012	-0.643
60/40 Port.	1.977	0.014	0.049	0.054	0.011	-0.562
Model (TCTP)						
1*	2.416	0.080	0.109	0.073	0.063	-0.608
2	2.143	0.070	0.105	0.062	0.053	-0.677
3	0.615	-0.035	0.110	-0.073	-0.053	-0.546
4	0.615	-0.033	0.122	-0.064	-0.056	-0.535
5	0.596	-0.037	0.111	-0.075	-0.056	-0.550
8	0.553	-0.042	0.121	-0.075	-0.064	-0.557
9	1.901	0.062	0.121	0.044	0.040	-0.557
10	0.615	-0.035	0.110	-0.073	-0.053	-0.546
11	0.651	-0.030	0.112	-0.065	-0.049	-0.502
12	1.210	0.022	0.112	-0.001	0.004	-0.453
13	2.132	0.070	0.114	0.058	0.051	-0.627
14	0.614	-0.035	0.109	-0.074	-0.053	-0.511
Model (TVTP)						
15	1.241	0.024	0.105	0.001	0.007	-0.474
16	1.140	0.018	0.118	-0.006	-0.003	-0.462
17*	4.634	0.137	0.121	0.131	0.115	-0.797
18	0.869	-0.005	0.120	-0.032	-0.026	-0.453
19	1.153	0.019	0.120	-0.004	-0.002	-0.562
20	1.279	0.028	0.124	0.006	0.005	-0.414
21	0.905	-0.000	0.128	-0.025	-0.025	-0.430

*Notes: The table displays the economic value of various HMM-based investment strategies. MVO Port., 100% MSCI World, and 60/40 Port. refer to the benchmarks portfolios defined in Section 5.2.2. Models 1-14 belong to the TCTP class of models, while models 15-21 belong to the TVTP class of models. R^{CUM} is the cumulative wealth of the strategy, with a starting investment of €1, $\hat{R}$ are the annualized average returns, $\hat{\sigma}$ are the annualized standard deviations, SR is the Sharpe ratio, CER is defined as the certainty equivalent return with γ set to 3. MaxDD is the maximum drawdown. More information on the different criteria can be found in Section 5.2.1. * here, indicates models that outperform the benchmarks in terms of risk-adjusted returns as measured by the SR and the CER.*

From Table 5.11, it is clear that model specification 17 corresponding to a **TVTP AR(1), switching variance with exogenous variable term spread model** delivers the highest (risk-adjusted) returns. Furthermore, model Specifications 1, 2, and 9 seem to perform at a similar level to the benchmarks in terms of (risk-adjusted) returns, while the other models show inferior performance. It should be noted, however, that the economic performance of these trading strategies is heavily influenced by the way transaction costs are taken into account.

Furthermore, a potential problem with the trading strategies and the weights they prescribe is that they often prescribe significant changes in weighting at every step. The institutional investors for whom this research is conducted might not be willing to or be able to adhere to such drastic swings due to regulations. This is a problem with many of the studies in the sector, as established by Nystrup et al. (2018). Figure 5.7 presents and compares the portfolio values of strategies based on model specifications 1, 2, 5, and 8 with the benchmark strategies. The initial capital for all investment strategies is €1000. The investment strategies based on model specifications 5 and 8 gradually erode the investor's wealth due to the fact that they prescribe an excessive number of regime switches between marketable securities and fixed-income securities.

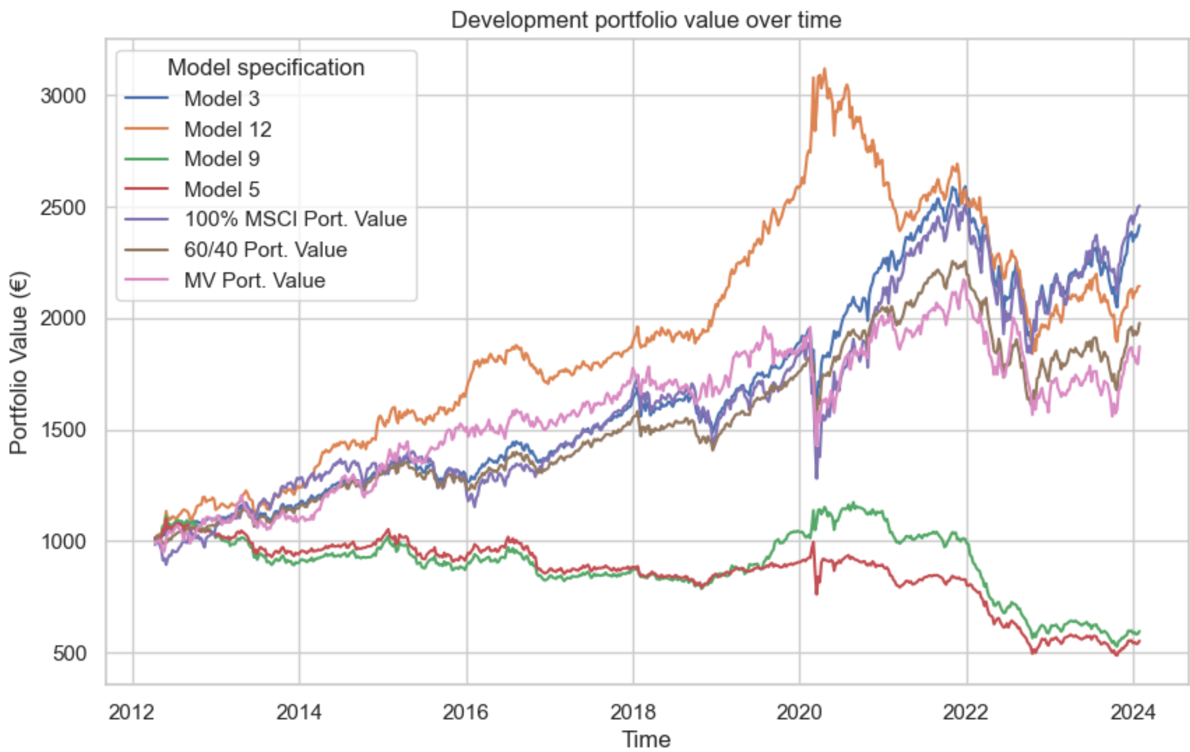

Figure 5.7: Development of cumulative economic value of several HMM-based models and the selected benchmarks for the time period from 2012-04-01 to 2024-02-01.

Note: The initial capital for all investment strategies is €1000.

Conclusion

The statistical accuracy and economic value of the HMM return forecasts are generally not significant from a statistical or economic point of view. The statistical performance of the HMM in terms of MSFE, RSME, and R^2 indicates that the considered models do not hold significant predictive power. The models' inability to accurately represent the amplitude of the actual return data is the root cause of this. The economic value generated by the return forecasts of certain HMMs is good in comparison to the various benchmark portfolios. Model specification 17 outperforms the benchmarks, including the 100% MSCI World portfolio. However, further research into the robustness of these results is needed before implementation. Additionally, problems may arise as a result of the strategies' frequent adjustments. Figure 5.8 depicts a scatter plot illustrating the relationship between the economic value generated by a model, quantified by R^{CUM}, and the statistical accuracy of the return forecasts, measured by R^2. The majority of the data points in the scatter plot suggest that there is a lack of a significant correlation between the two variables. The data point located at the extreme right of the graph corresponds to model specification 17. Based on this figure, it appears that this particular point is an outlier.

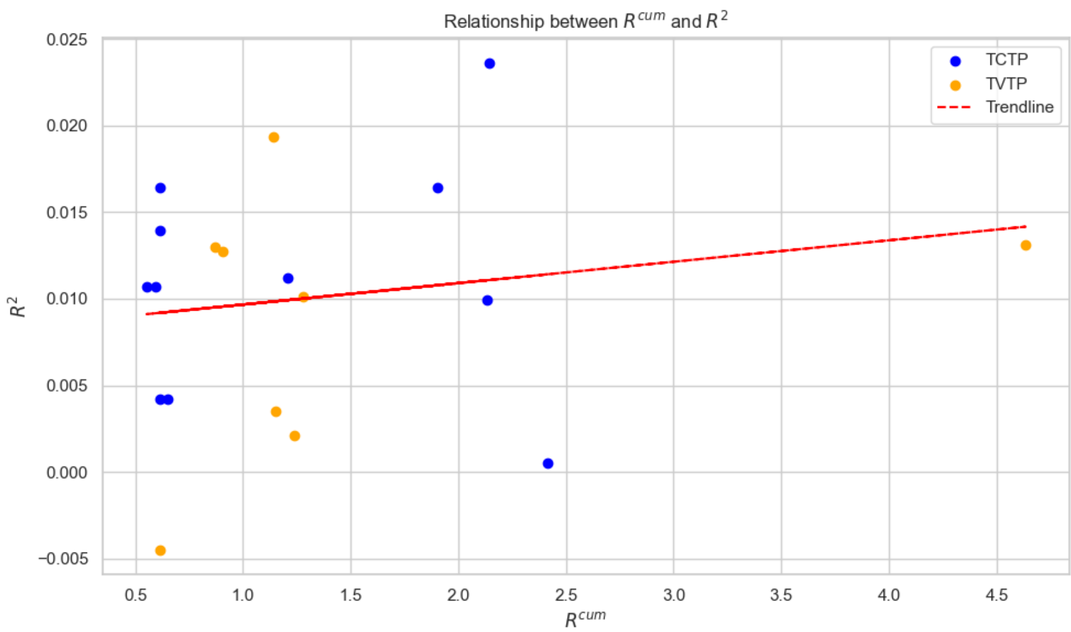

Figure 5.8: Scatter plot of economic value measured in cumulative returns and statistical accuracy measured in R^2

In conclusion, although the statistical accuracy may be lacking, the economic value derived from these return forecasts can be significant, which underscores the importance of considering both statistical and economic aspects when evaluating forecasting models.

Chapter 6

Conclusion

The starting point of this research was the need for Dutch pension funds to reevaluate their way of determining investment policies in light of the introduction of the new pension system. The main goal of this paper is to try to optimize the strategic asset allocation (SAA) of an institutional investor by incorporating tactical asset allocation (TAA). To achieve this goal, three research questions had to be answered first.

1. Which model is the most efficient in capturing stylized effects commonly observed in the return distributions of an institutional investor portfolio over time?

Through an extensive literature review, it became clear that hidden Markov models (HMMs) have attractive properties that allow them to capture many of the nonlinear patterns observed in the distribution of asset returns while also being tractable in the context of asset allocation. From further empirical research, it can be concluded that the time constant transition probabilities (TCTP) class of HMMs is the most effective in capturing non-linear patterns in the distribution of asset returns. All of the HMMs outperformed their standard autoregressive counterparts in terms of their BIC score. Moreover, the TCTP models outperformed their time varying transition probabilities (TVTP) counterparts. Additionally, within the class of TCTP HMMs, minimal differences were observed between specifications, indicating that the choice of exogenous variables is of limited significance in the regression analysis.

2. Are HMMs able to produce stock market regime forecasts and asset return forecasts that are significantly more accurate than random guessing?

From the research in this paper, one can conclude that the statistical accuracy of the out-of-sample stock market regime forecasts is significant. This conclusion is based on the statistically significant results produced by the HMMs considered in this paper. However, it should be mentioned that the regime-switch forecasting performance is weak, meaning that the models are not able to accurately predict stock market regime switches.

The statistical accuracy of the stock market return forecasts is far from significant; this holds for all of the models considered. Furthermore, no decisive answer can be given to the question that was raised earlier in Chapter 5 in regards to the efficacy of the testing procedure developed for regime forecasts.

3. Can HMM-based asset allocation models significantly outperform their standard linear-based counterparts in terms of risk-adjusted return in the context of an institutional investor like a Dutch pension fund?

Based on the evidence in this paper, it can be concluded that HMMs are able to outperform their linear counterparts in terms of risk-adjusted return. The risk-on/risk-off strategy in combination with the TCTP AR(1) class of HMMs delivers the most compelling evidence for this. Not only is this combination of strategy and model able to generate risk-adjusted returns even with the inclusion of transaction costs, but it also comes closest to the investment reality of an institutional investor and their governance structure. Furthermore, strategies that are based on return forecasts seem to perform relatively worse in comparison to strategies that are based on regime forecasts. Additionally, there seems to be a positive correlation between the statistical accuracy of the regime forecasts of a model and the economic value that it is able to generate.

By answering the research questions above, it has become evident that institutional investors like the clients of Sprenkels can increase the risk-adjusted returns on their portfolios by incorporating HMM-based TAA elements into their investment strategies. By considering both statistical and economic aspects, this study provides actionable insights for pension fund managers, enabling them to optimize asset allocation strategies and enhance risk-adjusted returns in the new pension system.

Chapter 7

Discussion

In addition to the conclusions that can be drawn from this study, there are also several interesting possibilities for further research. The depth of the literature surrounding asset allocation required me to be very selective in what to research. Because of this fact, there are still a lot of unknowns that, if uncovered, can improve the understanding of the potential impact of HMM-based investment strategies on the portfolios of institutional investors like Dutch pension funds and can lead to even greater increases in performance.

Firstly, the practical implications of implementing HMM-based asset allocation strategies require more in-depth research. As mentioned in Section 5.2.1 and Section 5.2.2 HMM-based strategies frequently require large shifts in wealth from one category to another. Assessing whether this is feasible in real-world settings, when considering factors such as governance structures, regulators, and other operational challenges, is perhaps the most important step. Evident from the research that has been done in this paper is that implementation of these HMM-based strategies requires significant changes in the governance structure of Dutch pension funds.

Secondly, it should be understood that while this study investigates a fair number of possible HMM specifications in depth, there are still a lot more model specifications and potential analyses to be completed. Furthermore, the potential gains that could be made with machine learning techniques or other more advanced econometric methods are also a fruitful area of further study. While this paper has focused on a more simple model that is easier to understand, it is undeniable that these techniques have potential. Implementation of machine learning techniques is especially interesting, since the renewed focus on machine learning as a result of the introduction of large language models (LLMs) such as ChatGPT has made machine learning accessible for everyone. Furthermore, the potential impact of including alternative metrics of performance like environmental, social, and governance (ESG) factors in the assessment of investment strategies could be interesting. These factors were excluded from this research on purpose, such that the only focus would be on risk-adjusted returns and feasibility. Future research could perhaps focus on explaining patterns in ESG criteria and stock market returns. Another crucial aspect that needs to be further researched before HMM-based models can be implemented in real-world settings is robustness. Improving the robustness of the research results would require assessing the performance of the HMMs under different markets, market conditions, and parameter settings. However, please note that even with robustness checks, it is essential to recognize that past performance is no guarantee for future success. The essence of this principle is perfectly captured in the widely recognized stock market adage, "Past performance is not indicative of future results."

Chapter 8

Bibliography

(2020). Federal reserve actions to support the flow of credit to households and businesses. https://www.federalreserve.gov/newsevents/pressreleases/monetary20200315b.htm.

(2023). Balans van pensioenfondsen kwartaal breukvrij. https://www.dnb.nl/statistieken/data-zoeken/#/details/balans-van-pensioenfondsen-kwartaal-breukvrij/dataset/0c3e2b24-63e3-4b61-8d78-87fc0b3b2e2d/resource/886821b7-4c50-42bc-8b43-260af7ddbbe7.

(2023a). The discount window and discount rate. https://www.federalreserve.gov/monetarypolicy/discountrate.htm.

(2023). The federal reserve - federal open market committee. https://www.federalreserve.gov/monetarypolicy/fomc.htm.

(2023b). Open market operations. https://www.federalreserve.gov/monetarypolicy/openmarket.htm.

Almgren, R. and Chriss, N. (2007). Optimal portfolios from ordering information. In *Forecasting Expected Returns in the Financial Markets*, pages 55–100. Elsevier.

Ang, A. and Bekaert, G. (2002). International asset allocation with regime shifts. *The review of financial studies*, 15(4):1137–1187.

Ang, A. and Timmermann, A. (2012). Regime changes and financial markets. *Annu. Rev. Financ. Econ.*, 4(1):313–337.

Anzuini, A., Lombardi, M. J., and Pagano, P. (2012). The impact of monetary policy shocks on commodity prices. *Bank of Italy Temi di Discussione Working Paper*, (851).

Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial analysts journal*, 48(5):28–43.

Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? *Journal of Banking & Finance*, 45:182–198.

Brocato, J. and Steed, S. (1998). Optimal asset allocation over the business cycle. *Financial Review*, 33(3):129–148.

Bulla, J. and Bulla, I. (2006). Stylized facts of financial time series and hidden semi-markov models. *Computational Statistics Data Analysis*, 51(4):2192–2209. Nonlinear Modelling and Financial Econometrics.

Çakmaklı, C. and van Dijk, D. (2016). Getting the most out of macroeconomic information for predicting excess stock returns. *International Journal of Forecasting*, 32(3):650–668.

Campbell, J. Y. and Viceira, L. M. (2002). *Strategic asset allocation: portfolio choice for long-term investors*. Clarendon Lectures in Economic.

Campion, M. A. and Campion, E. D. (2023). Machine learning applications to personnel selection: Current illustrations, lessons learned, and future research. *Personnel psychology*, 76(4):993–1009.

Carroll, R., Conlon, T., Cotter, J., and Salvador, E. (2017). Asset allocation with correlation: A composite trade-off. *European Journal of Operational Research*, 262(3):1164–1180.

Chen, S.-S. (2012). Consumer confidence and stock returns over market fluctuations. *Quantitative Finance*, 12(10):1585–1597.

Diebold, F. X., Lee, J.-H., Weinbach, G. C., et al. (1993). Regime switching with time-varying transition probabilities.

Dzikevičius, A. and Vetrov, J. (2012). Stock market analysis through business cycle approach. *Business: Theory and Practice*, 13(1):36–42.

Eychenne, K., Martinetti, S., and Roncalli, T. (2011). Strategic asset allocation. *Available at SSRN 2154021*.

Fabozzi, F. J., Gupta, F., and Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The journal of investing*, 11(3):7–22.

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

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

Gonzalez, L., Hoang, P., Massey, J., and Shi, J. (2006). Defining and dating bull and bear markets: Two centuries of evidence. *Multinational Finance Journal*, 10(1/2):81–116.

Guédon, Y. (2003). Estimating hidden semi-markov chains from discrete sequences. *Journal of Computational and Graphical Statistics*, 12(3):604–639.

Guidolin, M. and Hyde, S. (2012). Can var models capture regime shifts in asset returns? a long-horizon strategic asset allocation perspective. *Journal of Banking & Finance*, 36(3):695–716.

Guidolin, M. and Timmermann, A. (2006). An econometric model of nonlinear dynamics in the joint distribution of stock and bond returns. *Journal of Applied Econometrics*, 21(1):1–22.

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

Guidolin, M. and Timmermann, A. (2008). Size and value anomalies under regime shifts. *Journal of Financial Econometrics*, 6(1):1–48.

Guo, Q. (2022). Review of research on markowitz model in portfolios. In *2022 7th International Conference on Social Sciences and Economic Development (ICSSED 2022)*, pages 786–790. Atlantis Press.

Haase, F. and Neuenkirch, M. (2023). Predictability of bull and bear markets: A new look at forecasting stock market regimes (and returns) in the us. *International Journal of Forecasting*, 39(2):587–605.

Hamilton, J. (2016). Chapter 3 - macroeconomic regimes and regime shifts. volume 2 of *Handbook of Macroeconomics*, pages 163–201. Elsevier.

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*, pages 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of econometrics*, 45(1-2):39–70.

Hamilton, J. D. (2010). Regime switching models. In *Macroeconometrics and time series analysis*, pages 202–209. Springer.

Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

Haugh, M. (2016). Mean-variance optimization and the capm. https://www.columbia.edu/~mh2078/FoundationsFE/MeanVariance-CAPM.pdf.

Hayo, B., Kutan, A. M., and Neuenkirch, M. (2010). The impact of us central bank communication on european and pacific equity markets. *Economics Letters*, 108(2):172–174.

Hodnett, K., Hsieh, H.-H., et al. (2012). Capital market theories: Market efficiency versus investor prospects. *International Business & Economics Research Journal (IBER)*, 11(8):849–862.

James, G., Witten, D., Hastie, T., Tibshirani, R., and Taylor, J. (2023). *An introduction to statistical learning: With applications in python*. Springer Nature.

Jensen, G. R. and Mercer, J. M. (2003). New evidence on optimal asset allocation. *Financial Review*, 38(3):435–454.

Jondeau, E. and Rockinger, M. (2005). Conditional asset allocation under non-normality: how costly is the mean-variance criterion? *Available at SSRN 674424*.

Joyce, M., Lasaosa, A., Stevens, I., and Tong, M. (2010). The financial market impact of quantitative easing.

Kim, C.-J. (1994). Dynamic linear models with markov-switching. *Journal of econometrics*, 60(1-2):1–22.

Kim, C.-J., Nelson, C. R., et al. (1999). State-space models with regime switching: classical and gibbs-sampling approaches with applications. *MIT Press Books*, 1.

Kole, E. and Van Dijk, D. (2017). How to identify and forecast bull and bear markets? *Journal of Applied Econometrics*, 32(1):120–139.

Kondlo, M. (2016). A framework for regime identification and asset allocation. Master's thesis, University of Cape Town.

Liu, J. (2023). A new tail-based correlation measure and its application in global equity markets. *Journal of Financial Econometrics*, 21(3):959–987.

Lunde, A. and Timmermann, A. (2004). Duration dependence in stock prices: An analysis of bull and bear markets. *Journal of Business & Economic Statistics*, 22(3):253–273.

Ma, Y. and Zhang, J. (2016). Financial cycle, business cycle and monetary policy: evidence from four major economies. *International Journal of Finance & Economics*, 21(4):502–527.

Mandelbrot, B. B. and Mandelbrot, B. B. (1997). *The variation of certain speculative prices*. Springer.

Markowitz, H. (1952). The utility of wealth. *Journal of political Economy*, 60(2):151–158.

Markowitz, H. M. and Markowitz, H. M. (1967). *Portfolio selection: efficient diversification of investments*. J. Wiley.

Markowitz, H. M. and Todd, G. P. (2000). *Mean-variance analysis in portfolio choice and capital markets*, volume 66. John Wiley & Sons.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257.

Nystrup, P., Hansen, B. W., Larsen, H. O., Madsen, H., and Lindström, E. (2018). Dynamic allocation or diversification: A regime-based approach to multiple assets. *Journal of Portfolio Management*, 44(2):62–73.

Nystrup, P., Hansen, B. W., Madsen, H., and Lindström, E. (2015). Regime-based versus static asset allocation: Letting the data speak. *The Journal of Portfolio Management*, 42(1):103–109.

Organisation for Economic Co-operation and Development (OECD) (2024). Composite leading indicator (cli). Accessed on 27 February 2024.

Otranto, E. (2010). Asset allocation using flexible dynamic correlation models with regime switching. *Quantitative Finance*, 10(3):325–338.

Pesaran, M. H. and Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance*, 50(4):1201–1228.

Rom, B. M. and Ferguson, K. W. (1994). Post-modern portfolio theory comes of age. *Journal of investing*, 3(3):11–17.

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

Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting*, 1:135–196.

Turner, C. and Han, J. (2009). Portfolio optimization under time-varying economic regimes.

Walters, C. et al. (2014). The black-litterman model in detail. *Available at SSRN 1314585*.

Chapter 9

Appendix

9.1 Data frequency analysis

Table 9.1 contains an analysis of the effect that the data frequency has on the economic value of the regime forecasts. The models tested are defined in Table 5.1, and the time period considered in the analysis is from 2012-04-01 to 2024-02-01. From the table, it is clear that there is a downward trend in the return data when the data frequency is decreased.

Table 9.1: Economic value regime prediction: data frequency analysis

R^{CUM}	Daily	Weekly	Monthly	Quarterly
Model (TCTP)				
1	1.050	0.124	0.343	-0.258
2	2.001	1.597	1.257	0.900
3	1.954	1.791	1.415	1.511
4	1.922	1.813	1.433	1.312
5	1.954	1.794	1.415	1.011
6	1.922	1.813	1.433	1.312
7	2.750	1.794	1.541	1.011
8	2.770	1.819	1.642	1.313
9	2.651	1.819	1.643	0.912
10	2.610	1.791	1.543	0.983
11	1.556	1.803	1.441	0.912
12	1.577	1.875	1.443	1.012
13	1.713	1.756	1.341	1.014
14	1.811	1.804	1.354	1.016
Model (TVTP)				
15	1.55	1.341	1.035	0.633
16	1.713	1.404	1.035	0.634
17	1.758	1.308	1.036	0.522
18	1.566	1.426	1.437	0.788
19	1.543	1.437	1.335	0.788
20	1.666	1.494	1.035	0.801
21	1.905	1.382	1.038	0.699

9.2 Data information

Below an overview is given of all of the macroeconomic variables and asset categories indices is given, in terms of where and how the data was obtained.

Table 9.2: Macroeconomic variables and asset class tickers

Macroeconomic Feature	**Ticker**	**Data Source**
Leading Indicators OECD: GDP America	USALORSGPNOSTSAM	Fred
Leading Indicators OECD: GDP Europe	EA19LORSGPNOSTSAM	Fred
Spot Crude Oil Price	WTISPLC	Fred
Consumer Price Index	CPIAUCSL	Fred
Unemployment Rate	UNRATE	Fred
10-Year Treasury Rate	GS10	Fred
Housing Starts	HOUST	Fred
OECD Leading Indicator CLI	OECDLOLITOAASTSAM	Fred
Producer Price Index	PPIACO	Fred
Term Spread	T10Y2Y	Fred
Federal Reserve Rate	DFF	Fred
Asset category index name		
MSCI	^990100-USD-STRD	Yahoo Finance
MSCI EM	EEM	Yahoo Finance
13-m US Treasury bond	TLT	Yahoo Finance
13-m US Treasury bond (Long)	SHY	Yahoo Finance
High Yield Corporate debt (DM)	HYG	Yahoo Finance
EM High Yield Bond	EMHY	Yahoo Finance
Oil Index	UNG	Yahoo Finance
Gold Index	GDX	Yahoo Finance
Real Estate Index	VNQ	Yahoo Finance

9.3 Implementation

The algorithm will be implemented in Python; more specifically the package statsmodels 0.14.1 will be used to not only fit the Markov Regime Switching model to our data but also to obtain forecasts of future regimes and future asset returns. Below, a quick rundown of the methods used will be provided:

- The function *statsmodels.tsa.regimeswitching.markovregression.MarkovRegression* lets users define the variables and all other relevant model choices that influence the model and its outcomes.
- The function *statsmodels.tsa.regimeswitching.markovregression.MarkovRegression.fit* fits the model to the data by using the Hamilton filter. The expectation maximization algorithm is used to search the parameter space for an optimal outcome.

9.4 Influence of delays on economic value

One of the more problematic implications has to do with implementation delay of portfolio switches. Due to the structure and size of the average institutional investor, decisions on relocation of funds typically require several weeks of throughput time. Which is problematic, since gains in returns all but disappear when taking into account such delays. Evident from the tables in this section. The models tested are defined in Table 5.1, and the time period considered in the analysis is from 2012-04-01 to 2024-02-01.

Table 9.3: Regime forecasts: Economic value delay = 0.

Model	R^{cum}	$\hat{R}$	$\hat{\sigma}$	SR	CER	MaxDD
Model (TCTP)						
1	0.124	-0.039	0.044	-0.258	-0.040	-0.887
2	1.597	0.010	0.035	0.081	0.009	-0.483
3	1.791	0.012	0.041	0.086	0.011	-0.544
4	1.813	0.013	0.041	0.088	0.012	-0.543
5	1.794	0.012	0.041	0.086	0.011	-0.543
6	1.813	0.013	0.041	0.088	0.012	-0.543
7	1.794	0.012	0.041	0.086	0.011	-0.543
8	1.819	0.013	0.042	0.087	0.012	-0.557
9	1.819	0.013	0.042	0.087	0.012	-0.557
10	1.791	0.012	0.041	0.086	0.011	-0.544
11	1.803	0.012	0.041	0.087	0.012	-0.543
12	1.875	0.013	0.041	0.092	0.012	-0.558
13	1.756	0.012	0.041	0.083	0.011	-0.536
14	1.804	0.012	0.042	0.086	0.012	-0.550
Model (TVTP)						
15	1.341	0.006	0.035	0.051	0.006	-0.398
16	1.404	0.007	0.035	0.058	0.006	-0.412
17	1.308	0.006	0.036	0.046	0.005	-0.376
18	1.426	0.007	0.037	0.058	0.007	-0.459
19	1.437	0.008	0.035	0.061	0.007	-0.456
20	1.494	0.008	0.035	0.068	0.008	-0.458
21	1.382	0.007	0.038	0.052	0.006	-0.411

Table 9.4: Regime forecasts: Economic value delay = 1.

Model	R^{cum}	$\hat{R}$	$\hat{\sigma}$	SR	MaxDD	CER
Model (TCTP)						
1	0.006	-0.098	0.054	-0.521	-0.994	-0.099
2	0.048	-0.057	0.055	-0.302	-0.953	-0.059
3	1.046	0.002	0.045	0.013	-0.371	0.001
4	0.988	0.001	0.045	0.006	-0.374	-0.000
5	1.023	0.002	0.045	0.010	-0.374	0.001
6	0.988	0.001	0.045	0.006	-0.374	-0.000
7	1.023	0.002	0.045	0.010	-0.374	0.001
8	1.090	0.003	0.045	0.018	-0.358	0.002
9	1.090	0.003	0.045	0.018	-0.358	0.002
10	1.046	0.002	0.045	0.013	-0.371	0.001
11	0.990	0.001	0.045	0.006	-0.378	-0.000
12	0.985	0.001	0.046	0.005	-0.362	-0.000
13	1.109	0.003	0.044	0.020	-0.384	0.002
14	1.064	0.002	0.045	0.015	-0.381	0.001
Model (TVTP)						
15	0.050	-0.056	0.054	-0.299	-0.949	-0.058
16	0.060	-0.053	0.055	-0.275	-0.940	-0.054
17	0.078	-0.048	0.056	-0.243	-0.923	-0.049
18	0.098	-0.044	0.053	-0.239	-0.903	-0.045
19	0.117	-0.040	0.052	-0.222	-0.883	-0.041
20	0.072	-0.050	0.052	-0.276	-0.927	-0.051
21	0.047	-0.057	0.054	-0.307	-0.952	-0.059

Table 9.5: Regime forecasts: Economic value delay = 2.

Model	R^{cum}	$\hat{R}$	$\hat{\sigma}$	SR	MaxDD	CER
Model (TCTP)						
1	0.005	-0.100	0.055	-0.529	-0.995	-0.102
2	0.040	-0.061	0.055	-0.323	-0.960	-0.063
3	1.080	0.003	0.045	0.017	-0.387	0.002
4	0.967	0.000	0.046	0.003	-0.380	-0.001
5	1.005	0.001	0.046	0.008	-0.376	0.000
6	0.967	0.000	0.046	0.003	-0.380	-0.001
7	1.005	0.001	0.046	0.008	-0.376	0.000
8	1.160	0.004	0.045	0.026	-0.380	0.003
9	1.160	0.004	0.045	0.026	-0.380	0.003
10	1.080	0.003	0.045	0.017	-0.387	0.002
11	0.974	0.001	0.046	0.004	-0.389	-0.000
12	0.985	0.001	0.046	0.005	-0.364	-0.000
13	1.122	0.003	0.045	0.021	-0.376	0.002
14	1.088	0.003	0.045	0.018	-0.388	0.002
Model (TVTP)						
15	0.019	-0.075	0.055	-0.394	-0.981	-0.077
16	0.024	-0.071	0.056	-0.365	-0.976	-0.072
17	0.036	-0.063	0.057	-0.318	-0.964	-0.064
18	0.046	-0.058	0.054	-0.311	-0.953	-0.060
19	0.054	-0.055	0.052	-0.306	-0.945	-0.056
20	0.027	-0.069	0.053	-0.374	-0.972	-0.070
21	0.024	-0.071	0.054	-0.375	-0.976	-0.072