Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Bsc2 Econometrics and Economics

Market Timing using Macroeconomic Trees and Global Split Criteria

Thijs Ackers (525411)

Supervisor:	A Tetereva
Second Assessor:	EP O'Neill
Date final version:	2nd July 2023

Abstract

Factor investing, rooted in the capital asset pricing model (CAPM), has evolved to encompass many risk factors beyond the market factor. This paper explores the relationship between macroeconomic variables and optimal portfolio weights for multiple risk factors using decision trees and random forests. Decision trees are constructed using a global split criterion based on portfolio performance using sixty years of data on macroeconomic variables and factor returns, forming macroeconomic regimes. A subset resampling procedure is employed to enhance portfolio robustness, while random forests will be used to aggregate decision trees to improve out-of-sample performance. Empirical analysis reveals concrete investment recommendations that are aligned with previous research, such as investment in momentum factors depending on the past market return and in value stocks when the term spread is negative. The study demonstrates that while the tree portfolios outperform mean-variance portfolios pre-transaction costs, active trading erodes most Sharpe ratios. Furthermore, the trees can be used to extend existing models and hedge against specific factors, providing strong out-of-sample performance. Random forests exhibit slightly improve performance and significantly lower turnover. The importance of the term spread and market return in the models highlights the influence of the current and ancitipated state of the economy on portfolio construction. Furthermore, this paper shows that predicting and trading on macroeconomic signals can enhance portfolio performance.

1 Introduction

Factor investing has emerged as a prominent investment approach and finds its origin in the capital asset pricing model (CAPM) introduced by Sharpe (1964). Initially, the market factor was used as a single factor and the sensitivity of stocks with respect to this factor was measured by β (the market beta). After this work, many more risk factors have been found and documented, an overview of which can be found in Harvey, Liu and Zhu (2015). Well-known factors include the value and size factors considered by Fama and French (1992). This model was later extended to include a momentum factor as introduced by Carhart (1997), as well as profitability and investment factors, resulting in the influential five-factor model by Fama and French (2015).

In practice, asset managers employ multi-factor portfolio strategies, adjusting their exposure to various risk factors in response to market or macroeconomic signals. In order to benefit from these signals, a good understanding of the relationship between these signals and the performance of portfolios based on certain risk factors is of great importance. As emphasised by Foglia, Recchioni and Polinesi (2021) "this context gives rise to the need to investigate the correspondence between multi-factor portfolio strategies and the performance of the economy". This paper contributes to this by using decision trees and random forests to determine the optimal exposure to multiple factors as a function of macroeconomic variables.

This study uses decision trees and random forests to predict optimal portfolio weights for a wide range of asset classes and can be seen as an extension of the work by Cong, Feng, He and He (2021), who create panel trees to form cross-sections of stocks. More specifically, in this paper, decision trees will be constructed using the performance of a portfolio defined over all leaf nodes as a global split criterion. Trees of varying sizes and depths will be used to create macroeconomic regimes based on diverse macroeconomic variables, facilitating the prediction of optimal asset allocations within each of these regimes. The portfolio estimation procedure of subset resampling will be applied to obtain portfolios in order to enhance the robustness of portfolios formed on small samples in the deeper leaf nodes. Additionally, the aggregation of macroeconomic decision trees into random forests improves out-of-sample performance and allows for feature importance analysis. Furthermore, the trees are used to construct hedging portfolios for the S&P500.

The empirical analysis employs returns of a large selection of some well-known risk factors spanning the period 1963-2023. Several macroeconomic variables, including the term spread, GDP growth, and inflation are incorporated to grow the decision trees. The resulting trees offer some concrete investment recommendations for different economic regimes, aligning with findings in previous literature. The trees produce out-of-sample portfolios that outperform a mean-variance portfolio prior to transaction costs. However, actively following macroeconomic signals necessitates frequent trading, such that most Sharpe ratios are eroded by transaction costs. Additionally, the paper demonstrates that the trees can be used to extend any existing model, through benchmarking on that model. As such, the tree portfolios can be used to hedge against specific factors, such as the market. Building random forests from the trees yields portfolios with slightly superior performance and significantly reduced turnover, attributed to the averaging over multiple different portfolios. Notably, the term spread and the market return were found to be the most influential variables in the model, highlighting the consideration of the current and the anticipated state of the economy in constructing the portfolios.

As mentioned before, a proper understanding of the relationship between the macroeconomy and the returns of different risk factors is useful from a practical point of view. It enables asset managers to better capture economic signals and adjust their portfolios accordingly, thereby improving portfolio performance. From an academic perspective, this paper furthers the understanding of the application of decision trees to time-series data, and shows how an economics-guided global splitting criterion effectively handles time-dependence. Additionally, it provides practical insights into the concrete use of tree structures in portfolio construction and highlights the application of resampling methods to enhance the robustness of portfolios estimiated with limited samples.

This paper is structured as follows. In the second section, the existing literature on factor investing, models for economic regimes, and the technique of decision trees and random forests will be discussed. Then, in the third and fourth sections, the data and methodology will be presented. After this, in the fifth section, the main results will be presented, as well as a discussion of the optimal portfolio weights predicted by the models. The conclusion can be found in the sixth section of the paper.

2 Literature Review

Factor or style investing has been gaining attention since the introducton of the capital asset pricing model (CAPM) by Sharpe (1964). Since then, style investing has gained popularity from both academic and practical perspectives. Style is broadly defined by Kao and Shumaker (1999) as any system of classification by market segments that have distinguishing characteristics. In the literature, strong evidence has been found that style investing can yield returns in excess of those predicted by the CAPM. Recognising the limitations of the single-factor CAPM in return predictions, researchers have endeavored to develop multifactor models.

This effort has been spearheaded by Fama and French (1992) who focused their analysis on characteristic factors and divided the risk spectrum into three core components: size (as defined by market capitalisation), value (as defined by the book value to market value), and the market. This widely recognised Fama-French 3-factor model is often used as a base model and has been expanded with other factors such as momentum, dividend, liquidity, and quality, as demonstrated by studies including Carhart (1997) and Pastor and Stambaugh (2001)). The inclusion of these factors enhances the explanatory power of the models and results in improved asset pricing and investment performance. Notably, Jegadeesh and Titman (1993) documented the excess performance of a momentum-based portfolio, while Banz (1981) established a 'size effect', highlighting that smaller firms have had higher risk-adjusted returns than larger firms. These examples represent only a small subset of the diverse set of styles that have been established in the literature. Style decisions can have a significant impact on portfolio performance. As a result, timing portfolio strategies based on the past performance of different styles or factors has been an attractive approach for portfolio managers seeking to generate value.

2.1 Factor investing and macroeconomic variables

As mentioned earlier, timing strategies have been an attractive method for portfolio managers. Timing strategies can often be guided by macroeconomic variables that reflect the state of the economy. Consequently, a significant area of financial research aims to understand this relationship between macroeconomic variables and stock or factor returns. The rationale is that if one can establish the effect the state of the economy has on the returns of certain stocks or risk factors, one can adjust asset positions based on the changing state of the economy and achieve excess returns. There exist roughly two ways of linking macroeconomic variables and stock returns: (1) through the arbitrage pricing theory (APT), which employs a set of risk factors to explain asset returns, and (2) through the discounted cash flow model (Humpe & Macmillan, 2007).

Humpe and Macmillan (2007) use the discounted cash flow interpretation in a cointegration model to quantify the effects of various macroeconomic variables on stock prices. By modeling stock prices as the present value of expected future dividends, the authors demonstrate that any factor that influences cash flows or discount rates can impact stock prices. Their findings reveal that stock prices were influenced positively by industrial production and negatively by inflation and the interest rate during the period of investigation in the United States. Similarly, Chen, Roll and Ross (1986) model stock returns as a function of macroeconomic variables using regression analysis. Consistent with Humpe and Macmillan (2007), their results show that industrial production and, to a lesser extent, inflation are found to be significant in explaining stock returns. Additionally, changes in the risk premium and in the yield curve contribute to explaining excess returns. Furthermore, Jareño and Negrut (2015) establish a positive correlation of gross domestic product (GDP), industrial production, and inflation with the Dow Jones Index, while interest and unemployment rates exhibit a negative correlation. Although these studies confirm the existence of relationships, they employ relatively simple methods such as correlation and regression analysis. An increased understanding of these relationships can prove to be beneficial for portfolio management.

Tangjitprom (2012) provides a review of a number of studies on the effect of macroeconomic factors on stock returns. Roughly, the macroeconomic variables considered can be split up into those capturing general economic conditions (e.g., industrial production, employment), monetary policy (e.g., interest rates, yield curve, and money supply), price levels, and international activities (e.g., exchange rates). The most commonly used methodologies in this field of research encompass regression models, GARCH models, VAR models, Granger causality tests, and event studies. While the results are mixed, most studies support the notion of a relationship between macroeconomic variables and stock returns, for both short- and long-term horizons. Evidently, these studies predominantly rely on parametric models. However, considering the advancements in algorithms and non-parametric machine learning models since the publishing of this paper in 2012, there is potential for obtaining more conclusive results using more flexible and advanced methodologies.

As mentioned before, literature has established the ability of factor-based portfolios to outperform benchmarks. However, an essetnail element of factor investing is the cyclicality of factors. While factor indices demonstrate excess risk-adjusted returns over long periods of time, they can experience significant underperformance and cyclicality over short time horizons (Bender, Briand, Melas & Subramanian, 2013). To consistently outperform benchmarks, timing factor strategies based on macroeconomic variables that capture the state of the economy become crucial to guarantee the success of factor investing. Bender et al. (2013) document that factor performance has been cyclical, but that the periods of underperformance for different factors have not coincided. As such, this paves the way for diversification among different factors, giving rise to multi-factor portfolios. For instance, they find that a quality index performs well during recessions, whereas momentum and size-based strategies perform well during periods of expansion. Additionally, Ung and Luk (2016) analyse the performance of several well-known risk factors during different market and business cycle phases. Their findings indicate that factor strategies tend to be sensitive to market cycles, with growth and value factors performing well during bull markets, and low volatility and quality factors performing well during bear markets.

Timing strategies have been extensively discussed by a number of authors. Investors often diversify their investments across multiple factors to mitigate underperformance. However, as highlighted by Amenc, Esakia, Goltz and Luyten (2019), diversified factor portfolios often overlook the correlation between different factors and macroeconomic drivers, resulting in portfolios that are overly exposed to certain macroeconomic variables. Consequently, portfolios may fail to achieve balanced performance across different macroeconomic regimes. Taking into account these common correlations between factors and macroeconomic variables could lead to investment rules that provide multifactor portfolios that are diversified both across the factors and macroeconomic states, resulting in more robust and stable portfolio performance.

2.2 Factor performance and regime-switching models

A natural model to establish the varying performance of risk factors in different economic regimes is the Markov regime-switching model as documented by Hamilton (2005). These models allow for time-varying parameters across distinct economic states and define models to estimate transition probabilities between these regimes. Naturally, there exists a wide range of literature applying switching models to stock return predictions, but they are also valuable in the context of factor investing.

For instance, Chang (2009) employs a Markov regime-switching GARCH model to predict the dynamics of stock returns using macro factors such as the dividend yield, default premium, and interest rate. The author finds that the macro factors affect stock returns and that the degree of influence is time-varying depending on market volatility. Besides stock return predictions, regime-switching models are also applied to factor investing. For example, Ammann and Verhofen (2006) use a regime-switching model for the four-factor model by Carhart (1997). They identify a high-variance and a low-variance regime and find that value investing seems to be the most rational strategy in the former, while momentum investing performs better in the latter regime. Building upon this research, Gkatzilakis and Sivasubramanian (2014) establish the presence of regimes in smart beta indices. By forecasting these regimes and incorporating regime-dependent performance, they achieve higher Sharpe ratios through active portfolio intervention and combination of smart beta indices. Furthermore, Hammerschmid and Lohre (2018) use a Markov regime-switching model to estimate the current state of the economy and predict aggregate US stock returns using an identified macroeconomic regime factor.

Despite the intuitive appeal and extensive application of regime-switching models, the method has some limitations. The biggest drawback lies in model specification and estimation. A regime-switching model requires specification of the number of regimes in advance, as well as the specification of a model for and identification of transition probabilities and regime-specific parameters, making it less flexible. This can prove to be problematic when working with smaller samples, as it restricts the number of feasible regimes and the sample size for each regime. While several methods have been proposed to allow for more complex transition structures, increasing model complexity raises the risk of overfitting, especially in small samples. These drawbacks call for a more flexible and robust framework in order to better analyse the dependence between factor returns and macroeconomic variables and, subsequently, construct portfolios within different macroeconomic regimes.

2.3 Decision Trees

A popular intuitive non-parametric procedure that is often used within the area of machine learning is that of decision trees. The methodology of decision trees was introduced by (Breiman, Friedman, Stone & Olshen, 1984). In decision trees, a set of binary decision rules is used to split up the feature space (input variable space) into a set of rectangles (or leaf nodes) in such a way that, for example, the information gain or node purity is maximised. The main objective of the model is to make predictions for the different classes or leaf nodes resulting from the splits. Depending on whether the output variable is continuous or discrete, one speaks of classification or regression trees, hence the name Classification and Regression Trees (CART). As a result, with sufficient complexity, a tree structure can represent any continuous function or variable and capture non-linearities and feature interactions. As highlighted by Geurts, Irrthum and Wehenkel (2009) the main strength of tree methods is their interpretability. On the other hand, the main drawback is the high variance and hence instability, which often makes trees very sensitive to noise, causing the accuracy of single trees to fall behind that of other algorithms. Another drawback is that decision trees have a tendency to overfit to the data, resulting in poor generalisations to out-of-sample data.

The application of decision trees within finance is extensive. For example, Bryzgalova, Pelger and Zhu (2021) use decision trees to build cross-sections of asset returns that can serve as building blocks for new risk factors and test assets for asset pricing models. Singh (2022) applies several machine learning methods to predict stock prices, and finds that decision trees tend to perform similarly to more advanced models in-sample, but then fall behind out-of-sample, likely due to overfitting.

2.4 Random Forest

Random Forest (RF) is a non-parametric ensemble method that addresses the high variance of decision trees, as highlighted above, at the cost of slightly increased bias. Introduced by (Breiman, 2001), RF involves growing a large collection of decision trees, employing majority voting or averaging to generate predictions. It can be used for both data classification as well as regression tasks. RF combines the concept of 'bagging' (bootstrap aggregating) introduced by (Breiman, 1996), with random feature selection at each split. Bagging averages diverse models to reduce variance and enhance the model's performance. The random feature selection of RF reduces correlation among different trees and hence improves out-of-sample performance. Consequently, the technique possesses several desirable characteristics, such as speed, accuracy, and robustness. Furthermore, the degree of model tuning is primarily limited to one variable: the number of estimators or trees. Additionally, random forest facilitates the quantification of feature importance and the effect features have on the output variable, aiding in understanding the underlying effects at play. As a result, the technique has found extensive application in finance but also in other areas of research. For instance, a wide range of papers uses random forests to predict the direction of stock market price movements and construct portfolios accordingly (see for example, Khaidem, Saha and Dey (2016), and Basak, Kar, Saha, Khaidem and Dey (2019)). The literature is lacking when it comes to the application of random forests to macroeconomic data and factor returns, while the flexibility of random forests could be excellent at capturing these patterns in the data.

Decision and regression trees have gained popularity in machine learning due to their ability to accommodate nonlinear effects of predictors and capture interactions among them. In fact, trees are designed to group together observations based on some objective using predictor variables. The objective function often considers the degree of 'impurity' or dissimilarity among observations in different nodes, with splits designed to minimise this impurity locally. 'Locally' means that when considering a certain split, the model only considers data points available in the node that will be split. For independently and identically distributed (*i.i.d.*) data, the local split criteria are reasonable, as there is no interdependence between observations in different nodes (Saha, Basu & Datta, 2021). However, for dependent data, such as time series data, local splits ignore the overall dependence structure in the data. As such, the effect a split has on the overall performance of the tree is generally not taken into account. This approach usually results in overfitting, since fewer observations are available to calculate split criteria as the tree grows (Cong et al., 2021). Some papers try to deal with this issue. For example, the paper by Bryzgalova et al. (2021) applies a pruning algorithm with a global split criterion as a way to reduce overfitting of the decision trees.

Cong et al. (2021) employed a novel approach to optimise splits by using the performance of a portfolio encompassing all leaf nodes as a global split criterion. Their research focused on constructing panel trees capable of reducing dimensions in an imbalanced panel of stock returns and generating latent factors that have applications in asset pricing and investing. Notably, the decision trees developed by the authors demonstrated robust out-of-sample asset pricing and investment performance. Moreover, the authors extended the decision tree structure to incorporate interactions between macroeconomic variables and asset characteristics, enabling macroeconomic splits along the time-series dimension. Consequently, their model effectively captured regime-switching behavior across different macroeconomic states.

Considering the relatively limited data history spanning 39 years, Cong et al. (2021) restricted the analysis to a single macroeconomic split at the root of the tree. Their findings revealed that splitting the time series on market volatility and inflation improves investment performance both in and out-of-sample. However, the authors only allow for one time-series split and use an ad-hoc splitting rule in their methodology. This paper aims to enhance the existing methodology by incorporating the concept of a global splitting criterion to identify optimal splits for multiple macroeconomic variables. By doing so, this paper endeavors to uncover additional patterns between macroeconomic variables and factor performance, with the ultimate goal of deriving valuable investment rules.

2.5 Robust Portfolio Construction

As highlighted in the previous section Cong et al. (2021) chose to only allow for one macroeconomic split in order to preserve a substantial number of observations in each of the branches. Since the introduction of the mean-variance framework by Markowitz (1952) estimation of the model inputs, the first two moments of asset returns, has received much attention. The classical mean-variance portfolio often achieves poor out-of-sample performance due to the difficulty of estimating these input parameters. Kan and Zhou (2007) show that performance can be especially hurt when the length of the estimation period becomes small, and, specifically, that the realised out-of-sample portfolio losses could be unbounded when the sample size is smaller than the number of assets.

A frequently used heuristic to overcome this issue of parameter estimation uncertainty is that of resampling. A review of portfolio resampling methods, advantages, and drawbacks can be found in Scherer (2002). This study emphasises that resampling results in more diversified portfolios that can beat Markowitz portfolios, in the case of long-only constraints. Furthermore, it deals with extreme allocations (corner solutions), that are often mentioned as another drawback of the mean-variance framework. However, the author also mentions that it is unclear whether this result can be generalised to a different setting. In light of the success of ensemble methods in advancing the performance of existing algorithms, Shen and Wang (2017) expanded the approach of bootstrap aggregating to the area of portfolio resampling. They propose a new portfolio strategy particularly suited to situations with a large number of assets relative to the number of available data points. The authors resample subsets of the whole portfolio such that optimal subset portfolios can be formed using more accurate covariance estimates. By aggregating a number of optimal subset portfolios the authors obtain portfolios that outperform a wide range of benchmark portfolios on several datasets. The authors show that leveraging machine learning algorithms (bagging) into the problem of mean-variance portfolio optimisation can improve robustness and out-of-sample performance. The resampling approach can be a valuable addition to models in which one has to estimate robust portfolios using small samples. Consequently, using this procedure can be helpful when small samples are created after splitting a time-series into smaller sub-periods guided by macroeconomic variables.

3 Data

This research focuses on data from the United States because of data availability concerns. The period of analysis will range from July 1963 to April 2023 such that there is 60 years of monthly data. 70% of the data will be used to train the models (502 months) and the remaining 30% will be used to test model performance (216 months), giving rise to in-sample and out-of-sample observations.

In order to identify economic regimes in the United States economy, time series data are required on macroeconomic variables that capture the economic state. This paper uses a total of nine macroeconomic variables as features in the decision trees, all of which can be obtained from the Federal Reserve Economic Data (FRED) website.[1] These variables provide a good collection of variables that have been used in previous literature and are therefore believed to capture the most important macroeconomic economic effects. The variables are the following:[2]

- INF: inflation rate as measured by the monthly percentage change in the seasonally-adjusted U.S. Consumer Price index.
- GDP: GDP growth rate as measured by the quarter-over-quarter percentage growth in seasonally-adjusted U.S. real GDP. Monthly growth rates will be extrapolated from the quarterly growth rates.
- UNEM: Change in the unemployment rate as measured by the monthly percentage change in the seasonally-adjusted U.S. employment rate.
- IND: Growth rate of industrial production as measured by the monthly percentage change in the seasonally-adjusted U.S. Industrial Production Index.
- M1: Growth rate of the money stock is measured as the monthly percentage change in the seasonally-adjusted US M1 money stock.
- TS: Term Spread: as the term spread, the 10-year Treasury Bill rate with constant maturity minus the federal fund rate is used.
- PS: Pastor-Stambaugh Illiquidity obtained from the website of Robert F. Stambaugh.[3] This is a long-short portfolio that goes long in the decile of stocks most sensitive to liquidity shocks and goes short in the stocks least sensitive to liquidity shocks (Pastor & Stambaugh, 2001).
- Market (MKT): as a market factor the monthly return of the S&P500 is used.
- Risk-free rate (RFR): as the risk-free rate the monthly one-month Treasury bill rate is used obtained from Ibbotson Associates via the website of Kenneth French.

Table A1 provides an overview of some summary statistics of the macroeconomic variables and the correlation matrix of these variables can be found in Table A2. Furthermore, the NBER based recession indicator during the same period is used for some further analysis of the macroeconomic features.

Additionally, a set of investable assets is required and obtained from the website of Kenneth French.[4] All the details on how the portfolios were formed can be found on this website. Since one

[1] https://fred.stlouisfed.org/

[2] This list of variables is a combination of the variables used in Hammerschmid and Lohre (2018) (in which some of these are used to create economic regimes using Markov Regime-Switching models) and used in Cong et al. (2021) (in which some of these are used to grow decision trees on).

[3] https://finance.wharton.upenn.edu/ stambaug/

[4] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data$_{L}ibrary/f-f_factors.html$

of the goals of this paper is to see what kind of assets one should invest in, in different economic regimes, I chose a variety of 42 of the factor portfolios created by Fama and French of which monthly returns are available on the aforementioned website. This encompasses value-weighted portfolios of all NYSE, AMEX, and NASDAQ stocks related to several factors. More specifically, the analysis contains portfolios constructed using stocks in the top 30%, medium 40%, and bottom 30% of size, value, profitability, investment, earnings-to-price ratios, cashflow-to-price ratios, and dividend yields. Furthermore, three portfolios are included formed on stocks with negative earnings-to-price, cashflow-to-price, and dividend-to-price ratios. Next, the portfolios containing the top and bottom decile stocks based on the previous month's returns (short-term reversal), cumulative previous year's returns (momentum), and cumulative previous three-year returns (long-term reversal) are included to capture momentum effects. The last ten portfolios are formed using the top and bottom quintiles of all stocks based on accruals, the market beta, net share issues, the variance of daily returns, and the variance of the residuals from the Fama-French three-factor model. Together, these portfolios provide a wide range of investable portfolios of factors that have been shown to yield excess returns in some occasions. Hence, it will be interesting to see if one can use decision trees and random forests to uncover patterns between macroeconomic variables and optimal rate of investment in each of these factors.

4 Methodology

This paper extends and expands the methodology of Cong et al. (2021) by using panel trees and a global split criterion to split on macroeconomic variables. As such, decision trees will be used to create subsets of observations, that can be interpreted as macroeconomic regimes, and corresponding optimal mean-variance portfolios. Differently from normal decision trees, every possible split into subsets of observations will be evaluated on an overall split criterion, hence no longer requiring the data to be *i.i.d.* More specifically, the split, using a certain variable and split point, that maximizes the Sharpe ratio of the mean-variance portfolio using all data points will be used. This section will discuss the specifics of the methodology, as well as the performance measures that will be used to compare the models to some benchmark models.

4.1 CART versus Panel Trees

Cong et al. (2021) created a methodology that extends on the well-known Classification and Regression Trees (or CART for short), a decision tree that can be used both for the classification of observations as well as regression and prediction tasks (Breiman et al., 1984). CART can be used for asset pricing or investment objectives. More specifically, one can split an asset universe on several characteristic variables to divide assets into baskets and create portfolios (or factors) in each of these baskets. In this way, one can obtain a more diverse set of factors. There have been papers that show the superior performance of these kinds of factor models in asset pricing or constructing the Stochastic Discount Factor (see for example Bryzgalova et al. (2021)). However, when splitting a node, an algorithm based on CART only considers the observations within that node and computes a split criterion using only this information (for example, minimising the sum of squared errors). As a result of this local split criterion, the information contained in sibling nodes is not taken into account. Such an algorithm can be described as a recursive algorithm (depth-first) and this can result in overfitting and it does not incorporate any economic guidance or reasoning.

In order to solve this issue, Cong et al. (2021) came up with Panel Trees (panel because it can be applied to unbalanced panel data and it allows for both cross-sectional as well as time-series splits) that grow iteratively instead. At each iteration, a global split criterion is defined on all the leaf nodes, and that split is made that improves the performance metric over the entire dataset. Their analysis is focused on asset pricing using an extensive set of characteristic variables of the stocks, and their methodology allows them to create factors (mean-variance portfolios of the leaf portfolios) that perform well in terms of asset pricing.

4.2 Macroeconomic Panel Trees

This paper considers a universe of N factor portfolios over T time periods (months) and will split the entire set of months into sub-periods based on different macroeconomic variables, that should capture the state of the economy, using a global investment criterion. More specifically, assume a vector of K macroeconomic variables or features. Let the vector $X_t = (x_{t,1}, ...x_{t,K})$ denote these K macroeconomic variables at time $t = 1, ..., T$. At every iteration, a set of split rules will be defined,

$$\{\tilde{c}_{k,m}^{(j)} = (x_k, q_m) : j = 1, ..., S; k = 1, ..., K; m = 1, ..., M\} \tag{1}$$

where each split rule consists out of a macroeconomic variable and a certain value q_m, such that all observations t for which $x_{t,k} \leq q_m$ will enter the left split and the remaining observations will enter the right split. The cardinality of the set of split rules, S, depends on the number of variables that can be used to split on in a leaf and for a certain iteration, as well as the number of cutpoints considered per variable (the grid over which the model will search). These cutpoints can for example be defined by quantiles of the distribution of the macroeconomic variables corresponding to the observations within a certain node. At every iteration, the algorithm loops over all leaf nodes (the bottom nodes) and computes the values of the global split criterion for each of the split rules. The split rule (that is, the node, variable and split point) that achieves the minimum (or maximum) split criterion value, is chosen to split a leaf node on. As such, after the j-th split, the tree will generate $j + 1$ subsets of observations.

Each split candidate suggests a division of the time periods belonging to a certain leaf into two, possibly non-consecutive, sub-periods. In order to calculate the corresponding global split criterion, for each of the split rules one determines the optimal mean-variance portfolios for those two sub-periods. Subsection 4.3 will further expand on how exactly this portfolio is constructed. Essentially, one defines an optimal investment rule for that sub-period, where each of the sub-periods is defined by the values of several macroeconomic features.

To make it more concrete, imagine that at the first iteration, the root node (still containing all time periods), node 1, was split based on the monthly unemployment rate, where all periods with low unemployment rates were placed in the left node, node 2, and all periods with high unemployment rates were placed in the right node, node 3. These nodes become the new leaf nodes. In both leafs, one defines the optimal mean-variance portfolio computed with those observations. Hence, one forms the optimal investment rules for periods with low and high unemployment, respectively. The vectors containing the returns on these optimal portfolios are called the leaf basis portfolios and are denoted by $R_{2,t}^{(1)}$ and $R_{3,t}^{(1)}$, where the 1 indicates that these leaf portfolios have been formed after the first split. These portfolios only define investment rules within their corresponding sub-periods, such that the returns are equal to zero for all time periods t that are not in the specific leaf. The total Macroeconomic Panel Tree factor is defined as the union of these leaf basis portfolios. This implies that a fictional investor will fully alter the asset allocation to match the optimal allocation in a certain macroeconomic regime when a regime switch occurs. Mathematically, the tree factor $f_t^{(1)}$ after this first split is defined as

$$f_t^{(1)} = \bigcup_{i=2,3} R_{i,t}^{(1)} \tag{2}$$

or, generally, after the j-th split

$$f_t^{(j)} = \bigcup_{i=2,\ldots,j+1} R_{i,t}^{(j)} \tag{3}$$

where $R_{i,t}^{(j)}$ are all leaf basis portfolios after the j-th split.

Naturally, each split rule would have created different leaf portfolios and different tree factors as a result. The split rule is chosen in such a way that the negative squared Sharpe Ratio of the resulting tree factor ($f_t^{(j)}$ at iteration j) is minimised, or, similarly, when the positive squared Sharpe Ratio is maximised. This results in the following global split criterion

$$\Lambda(\tilde{c}_{k,m}) = -\mu_F^T \Sigma_F^{-1} \mu_F \tag{4}$$

where $F = f_t$ is the factor generated by the Macroeconomic Panel Tree for each of the split criteria, with a sample mean and covariance, μ_F and Σ_F. By constructing the split criterion in such a way, at each leaf the optimal portfolios of the other leaf nodes are taken into account. Therefore, splits will only happen if it results in improvements in the Sharpe ratio of the resulting overall portfolio following the investment rules given by the mean-variance portfolios within each macroeconomic regime. This procedure will be repeated until the splitting of the tree no longer results in improvements of the overall factor Sharpe ratio, or until the growth of the tree is limited by certain hyperparameters. The split criterion (4) also allows for the inclusion of other factors besides the factor created by the tree. In that case, the construction of the tree factor relies on improving the squared Sharpe ratio over certain benchmark factors. For this research, the portfolios have been limited to long-only portfolios. Furthermore, at every node, the cutpoints used as splitting points for every variable will be computed for the data available in that node.

Once a tree is grown it can be used to produce an out-of-sample portfolio. More specifically, by looking at the values of the macroeconomic variables in the previous period $t-1$ we can predict in which leaf of the tree a certain period t will be and invest accordingly. That is, adjust the asset allocation to meet the optimal asset allocation in the predicted leaf. Then, out-of-sample performance of this portfolio can be traced.

There are several hyperparameters that can be used to guide the tree-growing procedure. One can, for example, limit the growth by setting a maximum depth, or a minimum number of time periods in each leaf, or by simply limiting the number of leafs a tree can have. These and others will be discussed in more detail in Subsection 4.5. The pseudocode of the entire procedure can be found in Algorithm 1.

Algorithm 1 Macroeconomic tree growing procedure

Input: Asset returns $r_{i,t}$ and ranked macroeconomic variables $x_{i,t}$;
Output: Grow the tree from a root node, form leaf basis portfolios and the corresponding tree factor;

for $j \leftarrow 1$ to J **do**
  if current depth $\geq \mathrm{d}_{max}$ **then**
    return
  else
    Search the tree, and find all leaf nodes $\mathcal{N}$
    for each leaf node N in $\mathcal{N}$ **do**
      for each split rule candidate $c_{k,m,N}$ in $\mathcal{C}_N$ **do**
        Partition data temporally in N according to $c_{k,m,N}$.
        if Either left or right child of N does not satisfy minimal leaf size **then**
          $\mathcal{L}(c_{k,m,N}) = \infty$
        else
          Calculate leaf basis portfolios.
          Estimate factor using all leaf basis portfolios as in 3.
          Calculate the split criteria as in 4.
        end if
      end for
    end for
    Find the best leaf node and split rule that minimises the split criterion:

$$c_j = \underset{N \in \mathcal{N}, c_{k,m,N} \in \mathcal{C}_{\mathcal{N}}}{\operatorname{argmin}} \mathcal{L}(c_{k,m,N}) \tag{5}$$

    Split the node selected at the j-th split rule of the three c_j.
  end if
end for
return $\{R_t^{(}J), f_t^{(}J)\}$

4.2.1 Boosted Macroeconomic Trees

As mentioned in the previous section, the split criterion as stated in (4) can also be extended to include other factors. More specifically, one can define $F = [m_t, f_t]$, where m_t is a certain benchmark factor. In that case, the model will look specifically for a factor f_t that achieves improvements in the squared Sharpe ratio relative to the benchmark factor. Note that the objective of the Sharpe ratio refers to the mean-variance portfolio built on the benchmark model together with the generated factor. As such, the tree factors can be formed in such a way that they generate portfolios that augment any existing model or factor, improving investment performance accordingly.

In this study, the macroeconomic trees will be used to improve over the market factor, hence, with m_t as the S&P500 monthly returns. As explained by Cong et al. (2021), the generated factor can potentially provide a hedging portfolio for the market, with ideally a low correlation with the market factor. Hence, these market-benchmarked tree factors have the potential to form portfolios that achieve better performance than those formed by the standard trees. Naturally, the market variable will not be used to grow the trees, such that eight macroeconomic variables remain. The performance of equally-weighted (50-50) and mean-variance portfolios of the market and the factor produced by each of the market-benchmarked trees will be analysed.

4.3 Subset Resampling

As elaborated on in the literature review, it is well known that estimation errors in input parameters of the mean-variance model, the first two moments of the returns, might substantially hurt performance out-of-sample and in practice. This is especially an issue when the number of time periods is not much larger than the number of assets that is available to invest in (Shen & Wang, 2017). In the model used in this paper, when splitting the dataset across the time horizon, each new leaf will contain a smaller number of observations than the previous one, hence making accurate estimation of the optimal mean-variance weights a bigger challenge. Specifically, when going deeper into the tree, one might end up in a situation where the number of months in a leaf (T) is smaller than or similar to the number of assets available to invest in (N). As a result, the sample covariance matrix that is required to execute mean-variance portfolio optimization might no longer be positive (semi-)definite and hence no longer invertible. This makes it impossible to compute the optimal portfolio weights in that leaf. Furthermore, when T gets close to N, it hurts the robustness of the procedure.

In order to deal with this issue of dimensionality, Shen and Wang (2017) develop a method in which one resamples subsets of the original asset universe and finds the optimal mean-variance portfolio weights for that subset of assets. Because one takes a subsample of the assets, N is reduced and kept (substantially) smaller than the number of observations (T), which allows for a more accurate estimation of the input parameters. By aggregating a big number of these subset portfolios, one creates a well-diversified mean-variance portfolio of all assets. The authors call this aggregated portfolio the subset resampling portfolio (SSR). More specifically, for a large number of s times, one selects a subset of size b from the n available assets. Formally, let $I_j \subset \{1, ..., n\}$ denote the index set at iteration j with $|I_j| = b$ for $j = 1, ..., s$. Furthermore, let $\{R_k^{j,b}\}_{k=1}^{\tau}$ denote the return data for the subset of assets at iteration j, where τ denotes the number of time periods available for estimation (the number of months within each leaf). For each subset, one first calculates the $b \times b$ sample covariance matrix and the $b \times 1$ sample means. Then, one solves the mean-variance optimization problem and computes the optimal subset portfolio weights, $\hat{\omega}^{j,b}$, which is prolonged to an $n \times 1$ vector $\hat{\omega}^j$ by including zero weights for all assets not included in the subset I_j. The final portfolio weights are obtained by averaging over all s subset portfolio weights:

$$\hat{\omega} = s^{-1} \sum_{j=1}^{s} \hat{\omega^j} \tag{6}$$

SSR partially sacrifices benefits from diversification in exchange for smaller estimation risk.

In order to benefit from this mitigation of estimation errors, one needs to put $b < min(n, \tau)$. As Shen and Wang (2017) mention, there is a trade-off when it comes to choosing b: a larger subset allows for more benefits from diversification but requires estimation of a larger covariance matrix, thus resulting in bigger estimation errors. In this paper, the choice for b is guided by the minimum leaf size, because this parameter determines the minimum number of months one could find in a leaf. With regards to s, more iterations are naturally always preferred but come at the cost of a computational burden. In this paper, given the relatively small number of available assets, s will be set to 1,000 for the trees as discussed earlier, and 500 for the trees used for the random forest procedure. The pseudocode for this method can be found in Algorithm 2.

Algorithm 2 Subset Resampling Portfolio

Input: τ: number of periods for estimation; $\{R_k\}_{k=1}^{\tau}$ historical return data; n: number of assets; b: subset size; s: number of subsets;
Output: The vector of portfolio weights $\hat{\omega}$;

for $j \leftarrow 1$ to S **do**
  Randomly sample a set I_j of b indices from $\{1, ..., n\}$ without replacement;
  Select the return data as $\{R_k^{j,b}\}_{k=1}^{\tau}$;
  Compute the sample covariance matrix as $\hat{\Sigma}_{\tau}^{j,b}$;
  Compute the sample means of returns as $\hat{\mu}_{\tau}^{j,b}$;
  Compute the optimal subset portfolio weights $\hat{\omega}^{j,b}$ by solving a mean-variance optimization based on the estimated parameters $\hat{\Sigma}_{\tau}^{j,b}$ and $\hat{\mu}_{\tau}^{j,b}$;
  Construct the weights for the whole portfolio $\hat{\omega}^j$;
  for $i = 1, ..., n$ **do**
    $\hat{\omega}_i^j = \hat{\omega}_i^{j,b} \times \mathbb{1}\{i \in I_j\}$
  end for
end for
Aggregate the constructed portfolio weights based on s resamples as $\hat{\omega} = s^{-1} \sum_{j=1}^{s} \hat{\omega^j}$;

4.4 Random Macroeconomic Forest

A powerful procedure that is often applied to Classification and Regression Trees (CART) is that of random forests as introduced by Breiman (2001). The idea of random forests can be extended toward the Macroeconomic Panel Trees. Random forest is an ensemble approach that fits multiple trees and then combines the predictions of these trees to reduce prediction errors. The random component is twofold: (1) for each of the trees a random subset of the variables is chosen to grow the tree, and (2) bootstrapping is applied to obtain a random subsample from the complete sample that is used to train the tree. As such, the correlation between the trees is reduced. Due to this random component, in-sample overfitting tends to be reduced, and this improves out-of-sample performance. The random forest procedure will be applied to the Macroeconomic Panel Trees in order to see if out-of-sample portfolio performance can be improved relative to the fully grown trees and to get insights into variable importance.

More specifically, the data will be bootstrapped along the time-series dimension with replacement, such that each tree is grown on a random subsample of the available observations. Since bootstrapping occurs along the time-series dimension, the full cross-section of assets is preserved in each time period, which allows the model to capitalise on the low-serial correlation that is present between asset returns. Low-serial correlation is a key assumption of random forest. Each tree is grown on the full feature set for given hyperparameter values, but at each split point, one randomly draws three out of nine macroeconomic variables to consider as recommended by Breiman (1996). Finally, one independently grows a Macroeconomic Panel Tree on each of the bootstrap samples. A crucial hyperparameter for random forest is that of the number of estimators or the number of trees. Section 4.5 will further expand on how this parameter was set. Once a tree is grown, it can provide out-of-sample predictions for the optimal asset weights. These asset weights are averaged over all the trees in order to produce one big out-of-sample portfolio guided by predictions made by all trees.

4.4.1 Variable Importance

The random forest approach of growing a larger number of trees also allows for the quantification of variable importance. This will give more insight in how the trees are grown and which variables are of particular importance. Here, I follow a similar procedure as in Cong et al. (2021). Intuitively, the more often a variable is selected to split on, the more important this variable is in improving factor performance in terms of the Sharpe ratio. By keeping track of how many times each macroeconomic variable is used to split on as a ratio of the number of times that variable was considered for a split, one gets a measure for variable importance. Mathematically, for each macroeconomic feature z:

$$P(\text{select } z) = \frac{\text{Number of times } z \text{ is selected to split on}}{\text{Number of times } z \text{ is considered at a split}} \tag{7}$$

Another way to capture variable importance is by looking at what Cong et al. (2021) call a variable's treatment effect. In this study, the treatment effect of each variable is measured as the degree to which splitting a node on that variable reduces (increases) the (positive) squared Sharpe Ratio of the resulting factor. Hence, the characteristic importance of each macroeconomic feature z is defined as follows

$$\text{Characteristic Importance}(z) = \frac{1}{N_z} \sum_{j=1}^{N_z} \left[(-SR^2_{f_{j+1}}) - (-SR^2_{f_j}) \right]$$

where SR_{f_j} is the Sharpe Ratio of the factor f before split j, and N_z is the number of times a variable z was used to split on. Hence, this characteristic importance is defined as the average decrease in the squared Sharpe Ratio over all splits on a certain variable z.

4.5 Hyperparameters

In order to compare the performance of different models in the analysis of macroeconomic panel trees, several hyperparameters are considered. One such parameter is the number of nodes or splits in the tree, which influences the trade-off between the size of the leaf nodes and the number of leaf nodes. Increasing the number of leaf nodes results in smaller subsets of observations within each leaf, potentially leading to higher estimation uncertainty. This can be harmful to the out-of-sample performance of the mean-variance portfolios. To mitigate this effect, Subset

Resampling is employed. Conversely, in certain cases, allowing a leaf to be split further and making it contain a specific subset of observations may discover additional patterns and hence enhance out-of-sample performance, particularly when a different investment rule is required that will not be identified otherwise.

To examine the impact of different tree sizes, three types of models are compared. These models range from the least-grown to the furthest-grown trees, specifically: models with a maximum of ten splits and a minimum of 25 months per leaf, models with a maximum of fifteen splits and a minimum of twenty months per leaf, and models with twenty splits and a minimum of fifteen months per leaf. The primary reason for these different tree sizes and the primary focus will be to compare larger trees, with a greater number of leaf nodes, each containing a smaller number of observations, against smaller trees. Furthermore, each model is grown for two, three, and four cutpoints, where four cutpoints implies splits occur along the quintiles of the distribution.

When conducting the analysis of the Random Macroeconomic Forest, a crucial parameter that influences performance is the number of estimators or the number of trees that are grown. Again, there exists a trade-off. Using more trees might improve the accuracy of the model, but it also incurs higher computational demands. To monitor performance, the 'out-of-bag' Sharpe ratio is examined while training the random forest model. The out-of-bag subsample comprises all observations that were not used to train a given tree, so, all observations that were not a part of the bootstrap sample for that tree. After each tree is grown, the preceding trees are utilised to predict optimal portfolio weights for the out-of-bag sample, and its performance is evaluated by looking at the Sharpe ratio of that out-of-bag portfolio.

In this study, due to the high computational demands of growing the macroeconomic panel trees, the stopping criterion for tree growth is defined as the point when the out-of-bag Sharpe ratio falls below the average of the last ten trees, provided that at least fifty trees have been grown. Once this criterion is met, tree growth is halted, and the model is employed to predict out-of-sample weights and construct out-of-sample portfolios. When this criterion is not met, at most 500 trees will be grown due to computational limitations. Random forests are created for all combinations of the number of splits and cutpoints, as detailed earlier, to evaluate their respective performance.

4.6 Performance Measures

The trees will be compared both in and out-of-sample based on several performance measures. The performance of the macroeconomic trees and the random macroeconomic forest model will be compared to some well-known benchmark models: an equally-weighted factor portfolio, a value-weighted factor portfolio, and a mean-variance portfolio. For these benchmark models, the weights will be determined in-sample, and then a buy-and-hold strategy will be applied.

4.6.1 Sharpe Ratio and Jensen's α

Given that the trees are trained to deliver high Sharpe ratios, the first obvious performance measure is the Sharpe ratio:

$$\hat{SR}_k = \frac{\hat{\mu}_k}{\hat{\sigma}_k} \tag{9}$$

where $\hat{\mu}_k$ and $\hat{\sigma}_k$ are the sample mean and standard deviation of the time series of returns generated by the portfolio generated using strategy k.

Additionally, the different portfolios will be compared on the α they attain. α is a risk-adjusted measure that compares a portfolio's returns to that of a benchmark model. In this paper, the portfolios will be compared against the predictions of the Capital Asset Pricing Model (CAPM). More specifically, the α for a certain strategy k will be obtained by regressing the monthly portfolio returns against the monthly S&P500 returns.

Lastly, to find whether the models significantly outperform the benchmark models and whether certain parameters can significantly improve model performance, I will compare the Sharpe ratios of the different models using the Ledoit-Wolf Sharpe ratio test as described by Ledoit and Wolf (2008). This test utilises bootstrapping to create studentised confidence intervals for the difference between two Sharpe ratios and declares the Sharpe ratios as significantly different if zero is not contained within this interval.[5]

4.6.2 Turnover and Transaction Costs

The performance of the different trees will be compared based on the turnover of the portfolios, to get a sense of the amount of trading required to implement each of the investment strategies. This is an interesting measure because it is to be expected that the portfolios generated by the trees will have a larger turnover than the benchmark portfolios. The degree to which the turnover of the tree portfolios will exceed that of the benchmark models can give an indication as to how much of the returns will remain when transaction costs are included. The turnovers are computed as follows:

$$\text{Turnover} = \frac{1}{T-1}\sum_{t=1}^{T-1}\sum_{j=1}^{N}(|\hat{w}_{k,j,t+1} - \hat{w}_{k,j,t} \times R_{j,t}|) \tag{10}$$

in which $\hat{w}_{k,j,t}$ is the portfolio weight of asset j at time t when following strategy k, and $R_{j,t}$ is the return of asset j at time t. This formula adjusts for the return of each asset at time t because this alters the amount of rebalancing required at the end of time t to adjust the weights towards the optimal weights at time $t+1$.

In order to get an actual feeling of the Sharpe ratio after transaction costs, the return series will be adjusted for proportional transaction costs as introduced by DeMiguel, Garlappi and Uppal (2007). They assume a proportional transaction cost of c for trading in each of the assets. Consequently, one can denote the development of wealth for each strategy k as follows:

$$W_{k,t+1} = W_{k,t}(1 + R_{k,p})(1 - c \times \sum_{j=1}^{N}|\hat{w}_{k,j,t+1} - \hat{w}_{k,j,t} \times R_{j,t}|) \tag{11}$$

where the trade required in each asset at time t is defined as in 10, and $R_{k,p}$ is the return of

[5]For this test, a circular block bootstrap approach is used and the optimal data-dependent block size will be used.

the portfolio created by following strategy k. The return net of transaction costs is then given by:

$$R_{net,k,t} = \frac{W_{k,t+1}}{W_{k,t}} - 1 \tag{12}$$

This net-return series can then be used to compute the Sharpe Ratio net of transaction costs. Following DeMiguel et al. (2007), in this study a proportional transaction cost of 50 basis points per transaction will be assumed.[6]

5 Results

This section discusses the main results that were obtained. The first part will elaborately discuss the in-sample and out-of-sample performance of a set of trees that is grown on the original data and the full set of variables, as well as the market-boosted trees. Furthermore, it will contain a more in-depth discussion of the weights produced by the trees to get an understanding of the investment rules prescribed by the trees. The second part will discuss the outcomes of the random macroeconomic forest model including a discussion of the importance of different variables.

5.1 Optimal Macroeconomic Trees

Table 1 shows the various performance indicators for the fully grown trees for several combinations of hyperparameters as well as for the benchmark models. It must be noted that all trees were grown until no improvement in the Sharpe ratio of the obtained tree factor $f_t^{(j)}$ was achievable for any of the split rules considered at any node.

Panel A of Table 1 presents the in-sample performance of the models. One can see that almost all macroeconomic trees have superior performance to the benchmark models before transaction costs are considered. The models grown using two and four cutpoints obtain higher monthly average returns and significantly higher Sharpe ratios than all three benchmark models. The models grown using three cutpoints are outperformed by the mean-variance model in-sample. This is due to the fact that these models did not achieve improvements in the Sharpe ratio after one split, such that the growth of the trees was stopped. The slight differences in the numbers in the table for these models can be explained by the random element of Subset Resampling. Specifically, the first split took place based on the term spread, which is often interpreted as capturing expectations about economic growth, after which no further improvements of the Sharpe ratio were possible. This was caused by the specific cutpoint values since using different numbers of cutpoints does result more splits and consequently in higher average returns and Sharpe ratios.

One can see that the other trees attain Sharpe ratios that are 20 to 50% higher than that of the mean-variance portfolio in-sample. In general, Table 1 shows that bigger trees are not

[6] Just like emphasised in DeMiguel et al. (2007) it is important to realise that in the presence of transaction costs as defined here, it might not be optimal to implement the same strategy, since the model does not take these into account. Furthermore, since the assets are value-weighted portfolios of individual stocks, it is likely that the transaction costs will be higher than presented here. However, no good estimate for the transaction costs was available, hence, the best alternative was used.

necessarily required for better performance, as the highest in-sample Sharpe ratios are obtained by the two 'smallest' trees, namely those trees that are grown using two cutpoints and number of splits capped to ten and fifteen. This indicates that there is only a subset of the macroeconomic features that provides useful information, such that only the first splits result in substantially better performance, after which only gradual improvements are obtained. What is notable and expected is the higher turnover that comes with the more sophisticated tree models as compared to the benchmark models that require relatively little trading. When looking at the Sharpe ratio net of transaction costs computed on the net return series as in (12), it is confirmed that the high turnover of the tree portfolios completely erodes the Sharpe ratios. The table shows that no model significantly outperforms the mean-variance model in-sample, even though the three best performing models still attain a higher Sharpe ratio. This shows that market timing for factors comes at the cost of high turnover, such that the models are subject to high transaction costs. All models have significant alphas relative to the CAPM.

Panel B of Table 1 presents the performance of the models in the testing period. One can observe that the Sharpe ratios drop compared to the ones attained in-sample, which indicates some degree of overfitting. We can see similar patterns in performance as in-sample, namely that some of the macroeconomic trees do achieve higher Sharpe ratios than the benchmark models, with two models significantly outperform the best-performing benchmark model (the value-weighted model). This indicates that creating different investment rules based on different macroeconomic variables does provide some scope for better out-of-sample performance when transaction costs are ignored.

Similar to in-sample performance, turnover is higher for the portfolios based on the trees, and the Sharpe ratios drop substantially when transaction costs are taken into account. Notably, none of the tree models significantly outperforms the mean-variance portfolio. These results confirm once more that the weakness of the tree models is their high turnover, and that this can significantly worsen performance. Interestingly, one can see that the models that allow the largest number of splits (twenty) require lower turnover in as well as out-of-sample for all numbers of cutpoints. This could be because the distances between observations in leaf nodes closer together are smaller, such that the trading required when adjusting the asset allocation to that of a different adjacent leaf is smaller.

Furthermore, out-of-sample one can see that the 'bigger' trees, grown using the biggest number of cutpoints, perform best in terms of the Sharpe ratio. On the same note, the models that performed best in-sample (those with two cutpoints and the number of splits equal to ten and fifteen) perform worst in-sample and are even outperformed by some of the benchmark models. Notably, the models obtained using three cutpoints outperform the models using two cutpoints and the mean-variance portfolio, while in-sample this was the other way around. This is promising because these models require relatively little trading but still achieve higher mean returns and Sharpe ratios than some other tree-based portfolios and benchmark models, even though these models only contain two leaf nodes. This shows that by only considering signals from one macroeconomic feature, one can already improve out-of-sample performance. Similar to the in-sample results, all tree-based and benchmark models achieve significant alphas relative to the CAPM.

5.1.1 Analysis of a Tree

In order to get a first idea of the importance of the different macroeconomic variables, this section provides a brief overview of the variables that were used most often to split on. Out of all the splits in the trees presented in Table 1, 29.8% were performed on the term spread, followed by the general market (19.1%). A positive term spread is often seen as an indication of a positive view on future economic growth and the market return can also be seen as a variable capturing how well the economy is performing. Hence, the current state of the economy and the expected direction in which the economy is heading are the most important factors in determining optimal investment rules in this model, accounting for close to 50% of all splits. Interestingly, the actual GDP growth accounts for only 9% of the splits. This might be because this measure is backward-looking, and therefore less indicative of the current or future economic state. Less common were splits based on the growth of the U.S. industrial production and the Pastor-Stambaugh Illiquidity measure. A more extensive overview of the variable importance will be provided in the section that discusses the results of the random forest models.

Figure B3 contains the fully grown tree for four cutpoints and fifteen splits (the best-performing model out-of-sample). This figure can be used to better understand the tree-growing procedure and to deduce some first relationships between the variables and the factors. Following the general pattern as mentioned in the previous section, this tree splits the observations four times on the market and twice on the term spread. This roughly causes four baskets of observations in terms of the performance of the S&P500: the 'low' basket (with a monthly return in the previous month of less than -0.6%), the 'neutral' basket (return between -0.6% and 0.8%), the 'moderate' basket (return between 0.8% and 3.2%), and the 'high' basket (return higher than 3.2%). Notably, in the 'low' and 'neutral' baskets with low S&P500 returns, one can see relatively large weights for the value factor as well as the dividend yield factor. On the other hand, in all high-return regimes, investing in the short-term reversal factor is recommended. In the highest return regime, a relatively large position in the earnings-to-price factor is recommended.

5.1.2 Analysis of Portfolio Weights

In this subsection, the in-sample portfolio weights obtained from the fully grown trees will be analysed, since most of these models significantly outperform the benchmark models in terms of the Sharpe ratio. Therefore, a deeper look into the models' recommendation may provide useful investment insights. More specifically, I will analyse the average predicted weights by all of the nine models discussed in the previous subsection as a function of the macroeconomic variables, and uncover some patterns between the portfolio weights and the variables. Table A4 shows some summary statistics of the weights for all the assets, including the macroeconomic feature each portfolio weight has the highest correlation with. Figure B4 presents the correlations between the average weights and the macroeconomic variables. This table and figure give a first indication of which variables and portfolio weights are correlated, and hence in which assets one should optimally invest for each economic regimes. The figure shows that the portfolio weights are most correlated with the term spread (TS), the risk-free rate (RFR), and the market (MKT). This is expected, since TS and MKT were also used most often to split on, as explained in the

Table 1: Performance measures for several fully grown trees and benchmark models, in-sample as well as out-of-sample.

Panel A: In-sample (1963-2005)

Macroeconomic Trees						
Cutpoints	**Splits**	**Avg Return**	**SR**	**Turnover**	α	**Net-SR**
2	**10**	1.391	1.338**	0.414	1.356***	1.133
	15	1.600	1.544***	0.756	1.558***	1.175
	20	1.350	1.221*	0.335	1.336***	1.068
3	**10**	1.210	1.052	0.107	1.190***	1.005
	15	1.215	1.050	0.105	1.195***	1.005
	20	1.211	1.050	0.104	1.192***	1.005
4	**10**	1.336	1.220*	0.610	1.287***	0.938
	15	1.408	1.322***	0.915	1.368***	0.888
	20	1.316	1.163*	0.401	1.261***	0.986
Benchmark Models						
Equally-weighted		0.971	0.695	0.041	0.942***	0.681
Value-weighted		0.910	0.711	0.055	0.892***	0.690
Mean-variance		1.272	1.080	0.036	1.255***	1.064

Panel B: Out-of-sample (2005-2023)

Macroeconomic Trees						
Cutpoints	**Splits**	**Avg Return**	**SR**	**Turnover**	α	**Net-SR**
2	**10**	0.828	0.659	0.506	0.831**	0.456
	15	0.781	0.614	0.642	0.773*	0.362
	20	0.847	0.655	0.485	0.833**	0.467
3	**10**	0.916	0.728	0.085	0.940***	0.690
	15	0.916	0.724	0.084	0.938***	0.690
	20	0.914	0.720	0.083	0.937***	0.688
4	**10**	0.958	0.790*	0.741	0.959**	0.481
	15	0.974	0.823**	0.712	0.974***	0.522
	20	0.927	0.724	0.370	0.831**	0.578
Benchmark Models						
Equally-weighted		0.871	0.601	0.041	0.873**	0.587
Value-weighted		0.865	0.673	0.056	0.875***	0.652
Mean-variance		0.836	0.660	0.037	0.856**	0.646

Note. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$. Net-SR is defined as the Sharpe Ratio net of transaction costs (based on the return series in Equation 12). The Sharpe ratios are compared against the mean-variance portfolio using the Ledoit-Wolf Sharpe ratio test. All trees were grown with $s = 1,000$ and $b = 10$ for subset resampling. The growing of the trees was stopped when no gain in Sharpe ratio was achievable for any split rule. For the trees allowing 10, 15, and 20 splits, the minimum leaf size was set at 25, 20, and 15, respectively.

previous section. Some of the patterns between the variables and assets with the strongest correlations will be investigated.

Figure B5 shows the portfolio weights of several factors as a function of the term spread and the market return. As aforementioned, an inversion, or a negative term spread, signals investor pessimism regarding the economic outlook. Estrella and Mishkin (1996) were among the first to find that the US term spread is a strong predictor for economic recessions two or more quarters into the future. Consequently, the trees take up this strong signal of inversion and use it to guide investors towards or from certain assets to protect them from losses in a possible near recession. For example, Figure B5a shows that the model advises a large weight (more than 10%) in high value stocks when the term spread is small or negative. Athanassakos (2009) shows Canadian evidence that the value premium persists during bear markets in the period 1985-2005, such that investing in value stocks can be a save bet during recessions or when one expects a recession. Additionally, Figure B5e indicates that the models show a similar pattern for investment in the low investment factor (yearly change in capital investments) as a function of the term spread. The investment anomaly (namely that firms with high capital investments will see decreases in benchmarked stock returns) has been the subject of various research (see Titman, Wei and Xie (2004)), but there is no clear consensus on its cause. Prombutr, Phengpis and Zhang (2012) show that the investment anomaly can be explained by a Fama-French 3 model with factor loadings linked to the default spread as a business cycle variable. This indicates that there is some relation between the business cycle and the investment anomaly, however the literature does not agree on the direction of this relation. This model recommends investing in low-investment assets when recessions are expected. On the other hand, investment in the small size factor is advised for periods with high term spreads and hence a positive economic outlook (Figure B5c). This is in line with evidence found by Kim and Burnie (2002), who find that the small firm effect is most present in expansionary periods. The model advises investing in small firms when expansion is expected in order to possibly capitalise on the small firm effect.

The trees also make some recommendations for market regimes as defined by the past month S&P500 return. In periods with high previous month market returns, the model advises to increase positions in high momentum stocks and low long-term reversal stocks (Figures B5b and B5d). That is, the trees recommend investing in stocks that had high returns over the past 2 to 12 months, but low returns during the past 13 to 60 months. De Bondt and Thaler (1985) show that investors tend to overreact and heavily load on past winners. This causes markets to overreact in the short term, after which a market correction occurs in which past 'losers' experience large returns. By investing in high momentum stocks and low long-term reversal stocks when the market is performing well, one benefits from the overreaction in the short-term, while also gaining from the market correction. Hence, the trees recommend the investor to trade on the momentum and reversal factors during the high market return regime.

For investing in low market return regimes, the trees recommend increasing the position in high dividend yield stocks as shown in Figure B5f. Stocks that pay high dividends are often thought of as safe havens during bear markets, since investors believe that these companies have more stable financial results, and the dividend itself can protect investors from downside risk. Benartzi, Michaely and Thaler (1997) confirm this and show that firms that have increased dividends experience significant positive excess returns for a prolonged period. Consequently, investing in high dividend stocks when market returns are negative, may be a way to limit risks during a recession period. Concluding, the trees provide some concrete investment recommendations that are in line with findings in previous literature. Following these rules during different market regimes and periods with different economic outlooks can improve Sharpe ratios relative to benchmark models prior to transaction costs.

5.1.3 Boosted Macroeconomic Trees

Table 2 provides an overview of the Sharpe ratios including and excluding transaction costs for the 50-50 and the mean-variance portfolio of the market and the factors produced by the market-benchmarked trees. Tables A5 and A6 present the complete results for the portfolios, including mean returns, turnover and alphas. Figures B6 and B7 show the annualised returns of the market and the tree factors for the market-benchmarked model using 4 cutpoints and 20 leafs, in-sample and out-of-sample, respectively. Figure B6 shows that the tree factor has only positive returns in-sample, except for one year. Furthermore, it shows the ability of the factor to provide a hedge against the market, providing positive returns in all years were the market was down. Out-of-sample the factor shows a similar pattern, except for the year of the financial crisis (2008). Consequently, boosted trees can be used to benchmark on any existing model, and provide a factor that captures the macro state of the economy that can be used as a hedge against these models.

Table 2 shows that the portfolios perform well in-sample. Almost all portfolios (50-50 as well as a mean-variance portfolio) significantly outperform the mean-variance portfolio formed on all assets, reaching Sharpe ratios as high as 1.887. One can see that the models that consider the largest number of cutpoints perform best in-sample (4-10, 4-15, and 4-20). This gives an indication that the finest grid search over the macroeconomic variables results in the best portfolio performance in-sample as measured by the Sharpe ratio. This implies that for some variables looking at a small subset of the observations can capture valuable information that results in better performance. Furthermore, the mean-variance portfolios of the market and the factor produced by the market-boosted trees achieve higher Sharpe ratios than the equally weighted portfolios in-sample. However, out-of-sample this picture is reversed, with the 50-50 portfolio achieving higher Sharpe ratios for all models. This is most likely due to the parameter estimation uncertainty for the mean-variance model, resulting in worse out-of-sample performance. Out-of-sample, some models significantly outperform the mean-variance portfolio in terms of Sharpe ratio. Again, the models that use four cutpoints achieve the highest Sharpe ratios, with the model using four cutpoints and ten leafs attaining a Sharpe ratio larger than one.

Similarly to the non-boosted trees discussed in the previous section, the higher Sharpe ratios come at the cost of higher turnover. Tables A5 and A6 show that turnover can reach 0.9 for the mean-variance portfolios and 0.6 for the 50-50 portfolios. The turnover of the models is on average lower for the 50-50 portfolio than for the mean-variance portfolio, because holding the market results in lower turnover and all mean-variance portfolios held a position in the tree factors larger than half, hence requiring more turnover. However, opposite to the results presented in Table 1, some of the mean-variance portfolios produce Sharpe ratios net of transactions

Table 2: The Sharpe ratios before and after transaction costs for an equally-weighted and mean-variance portfolio of the market and the factor created by the market-boosted trees, for a variety of macroeconomic trees.

	50-50 Portfolio				Mean-Variance Portfolio			
	In-Sample		Out-of-Sample		In-Sample		Out-of-Sample	
Model	**SR**	**Net-SR**	**SR**	**Net-SR**	**SR**	**Net-SR**	**SR**	**Net-SR**
2 - 10	1.396**	1.180	0.838	0.615	1.472***	1.213	0.770	0.499
2 - 15	1.275	1.170	0.984*	0.840	1.314***	1.188	0.966**	0.797
2 - 20	1.447**	1.244	0.866	0.621	1.563***	1.311**	0.781	0.480
3 - 10	1.320**	1.164	0.901	0.728	1.359***	1.175	0.863	0.666
3 - 15	1.322**	1.169	0.903	0.729	1.362***	1.180	0.864	0.666
3 - 20	1.291*	1.191	0.932	0.805	1.328***	1.209*	0.896*	0.755
4 - 10	1.460**	1.185	1.041**	0.817	1.591***	1.220	1.002***	0.722
4 - 15	1.526***	1.155	0.959*	0.603	1.661***	1.170	0.880**	0.446
4 - 20	1.650***	1.240	0.981*	0.607	1.887***	1.298**	0.870*	0.404
MVP	1.080	1.064	0.660	0.646	1.080	1.064	0.660	0.646

Note. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. The Sharpe ratios are compared against the mean-variance portfolio in terms of the Sharpe ratio using the Ledoit-Wolf Sharpe ratio test. Net-SR is defined as the Sharpe Ratio net of transaction costs (based on the return series in Equation 12), under the assumption that holding the market is costless. MVP is the mean-variance portfolio of all assets. All trees were grown with $s = 500$ and $b = 10$ for subset resampling. The model names are defined as follows: number of cutpoints - maximum number of leafs.

that significantly outperform the benchmark mean-variance portfolio. Net of transaction costs, none of the models produce out-of-sample Sharpe ratios that are significantly higher than the benchmark model. However, one can see that the almost all net Sharpe ratios are higher than those in Table 1. This is partly due to lower turnover, but also due to the fact that the boosted trees produce less volatile portfolios.

5.2 Random Macroeconomic Forest

Table 3 shows the results for the portfolios obtained from the random forest models, in which the optimal weight predictions for a large number of trees were aggregated into one portfolio. Not all models had reached the stopping point where no improvements in the out-of-bag Sharpe ratio were possible, indicating that there were still performance improvements possible if there were no computational limitations.

When comparing the performance of the random forest portfolios in Table 3 to that of the fully grown tree portfolios in Table 1, a couple of developments stand out. First, one can see that the in-sample performance of the random forest portfolios has fallen below that of the tree portfolios, both in terms of average monthly returns as well as Sharpe ratios, which indicates a reduced degree of overfitting of the random forest portfolios. Now, most portfolios perform worse than the mean-variance portfolio in-sample both before and after transaction costs. Second, the random forest portfolios show a drastic decrease in turnover compared to the tree portfolios. Apparently, averaging over the weights predicted by many different trees reduces turnover. Consequently, the difference between the original Sharpe ratio and the Sharpe ratio net of transaction costs is smaller. All models still produce very significant excess returns over the CAPM predictions.

Out-of-sample the Sharpe ratios before transaction costs have increased relative to the fully grown tree portfolios in Table 1. Furthermore, similarly to the in-sample results, the portfolio turnovers have decreased for almost all models. Specifically, the average turnover dropped by almost 60% compared to the fully grown trees. Hence, applying the bagging approach to portfolio weights, reduces the volatility and turnover of the weights. Despite the slightly higher Sharpe ratios and the lower turnover, the Sharpe ratios net of transaction costs are not significantly higher than that of the mean-variance benchmark model. Hence, as expected, the portfolios formed by random forests do show slightly increased out-of-sample and drastically decreased turnover, but it is still not enough to significantly outperform the benchmark model.

5.2.1 Variable Importance

Figure 1 shows the selection probability for each of the macroeconomic features as explained in Equation (7), that is the probability that a variable was chosen to split on given that it was considered at that split. The black line is drawn at probability $\frac{1}{3}$, since this is the probability one would expect if one would randomly draw a feature at every split. Any variable that has a selection probability higher than this value is therefore seen as 'valuable' by the model. Here, one can see a similar pattern to the individual trees as mentioned in section 5.1.1. Namely, the term spread is chosen most often by the model to split on, with a selection probability of almost 50%. Additionally, the market and the risk-free-rate are selected close to 40% of the time. All the other features are chosen less than one would expect on a random basis.

These selection probabilities are partly in line with the treatment effects as shown in Figure 2. This figure presents the average decrease in the negative squared Sharpe ratio of the tree factor after a certain variable was chosen to split on, as explained in Equation 8. The term spread seems to be the most important variable of the model, resulting in the largest increase in the squared Sharpe ratio after each split and, consequently, the variable is the most important variable in forming the leaf nodes. The market variable has the second-highest treatment effect. However, interestingly, the risk-free-rate, which was the second feature in terms of selection probability, has almost the lowest average treatment effect in-sample.

Research mentioned earlier has shown that the term spread has a very high predictive power for recessions (see Estrella and Mishkin (1996)). This existence of such a relationship is also implied in the data used for this paper (see Figure B8a). Moreover, other research shows that the term spread also has high predictive power for other variables, such as GDP, industrial production, and the inflation rate (Kuosmanen & Vataja, 2011). Hence, one might also capture the effects of these variables by only considering the term spread. Table A2 shows that indeed the term spread indeed has the on average highest correlations with the other variables. This could provide an explanation as to why the term spread in the most dominant variable in this model. In general, the variable importance shows that the model mostly picks up information about the current and anticipated state of the economy (market and the term spread) and recommends investments in several types of assets accordingly. Figure B8 shows that the three most important variables in the model are all significantly correlated with the NBER recession indicator. This indicates that the presence of a recession seems to be the most important force

Table 3: Performance measures for the portfolios obtained with random forest on the macroeconomic trees, in-sample as well as out-of-sample.

Panel A: In-sample (1963-2005)

Macroeconomic Trees Cutpoints	Splits	Avg Return	SR	Turnover	α	Net-SR
2	**10**	1.110	0.941	0.094	1.131***	0.901
	15	1.149	0.971	0.123	1.126***	0.924
	20	1.232	1.122	0.098	1.271***	1.079
3	**10**	1.152	0.979	0.089	1.129***	0.942
	15	1.062	0.985	0.075	1.041***	0.950
	20	1.374	1.289	0.096	1.399***	1.244
4	**10**	1.148	0.960	0.097	1.123***	0.920
	15	1.062	0.985	0.075	1.108***	0.920
	20	1.183	0.986	0.105	1.155***	0.943
Benchmark Models						
Equally-weighted		0.971	0.695	0.041	0.942***	0.681
Value-weighted		0.910	0.711	0.055	0.892***	0.690
Mean-variance		1.272	1.080	0.036	1.255***	1.064

Panel B: Out-of-sample (2005-2023)

Macroeconomic Trees Cutpoints	Splits	Avg Return	SR	Turnover	α	Net-SR
2	**10**	0.878	0.745	0.166	0.882**	0.675
	15	0.907	0.717	0.163	0.917**	0.653
	20	0.900	0.728	0.165	0.912**	0.595
3	**10**	0.894	0.771**	0.164	0.910**	0.700
	15	0.908	0.811*	0.154	0.946***	0.742
	20	0.926	0.739	0.154	0.963***	0.679
4	**10**	0.969	0.800*	0.194	0.969***	0.718
	15	0.926	0.746	0.186	0.852***	0.672
	20	0.908	0.739	0.201	0.927***	0.658
Benchmark Models						
Equally-weighted		0.871	0.601	0.041	0.873**	0.587
Value-weighted		0.865	0.673	0.056	0.875***	0.652
Mean-variance		0.836	0.660	0.037	0.856**	0.646

Note. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$. Net-SR is defined as the Sharpe Ratio net of transaction costs (based on the return series in Equation 12). The Sharpe ratios are compared against the mean-variance benchmark model using the Ledoit-Wolf Sharpe ratio test. All trees were grown with $s = 500$ and $b = 10$ for subset resampling. The growing of the trees was stopped when no gain in Sharpe ratio was achievable for any split rule. For the trees allowing 10, 15, and 20 splits, the minimum leaf size was set at 25, 20, and 15, respectively.

behind the choices made by the model.

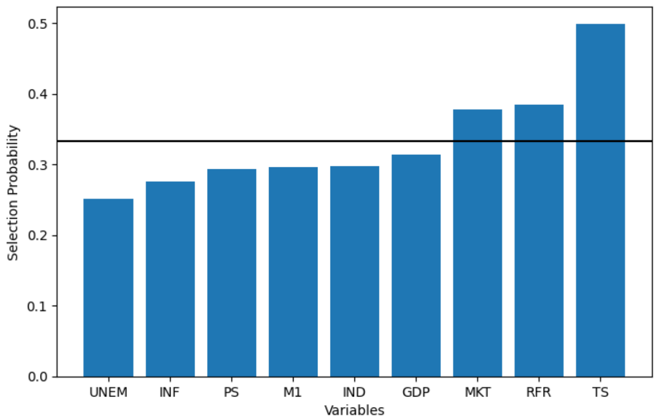


Figure 1: In-sample variable importance as measured by the selection probability in the random forest averaged over all models. The abbreviations of the variables can be found in Table A1.

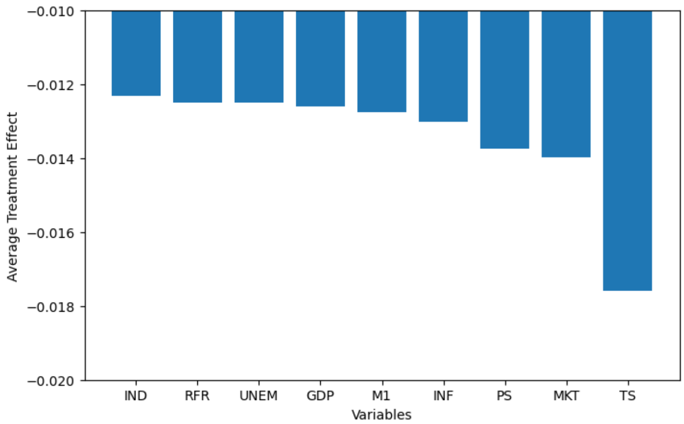


Figure 2: In-sample feature effect as measured by the decrease in negative squared Sharpe ratio after inclusion of this variable in the random forest averaged over all models. The abbreviations of the variables can be found in Table A1.

6 Conclusion

This paper created investment decision trees with macroeconomic features using a global split criterion that optimises the Sharpe ratio of a factor defined over all leaf nodes. It did so using 60 years of monthly data and nine macroeconomic variables, such that the decision trees split along the time-series dimension and hence create economic regimes. These regimes were created in such a way that each split results in the largest possible increase of Sharpe ratio of the tree factor defined as the concatenation of mean-variance portfolios defined over all the leafs. Furthermore, the assets used to create the portfolios form a wide range of well-known factor portfolios, such that the trees can help establish optimal factor investment rules within certain economic regimes. The portfolios were formed using subset resampling in order to enhance portfolio robustness. Performance measures such as the Sharpe ratio, turnover, and a transaction cost-adjusted Sharpe ratio were used to compare the portfolios against benchmark models, including mean-variance portfolios.

The results reveal that the macroeconomic trees advise portfolios with high turnover, resulting in significantly higher Sharpe ratios before accounting for transaction costs. However, due to the high turnover, most of the Sharpe ratios are eroded after transaction costs are taken into account. Notably, the trees provide some concrete investment rules that are in line with previous literature. The trees split most often on the term spread and the monthly market return. The models recommend investing in high value stocks when the term spread is negative and in small size stocks when the term spread is positive. Additionally, the models advise investment in the momentum factor and long-term reversal stocks to capitalise on short-term market overreactions and long-term market corrections.

Furthermore, the trees can be benchmarked against a specific factor or model, such that the trees will search for specific improvements in the Sharpe ratio over that factor or model. These trees are referred to as boosted macroeconomic trees. Using the market as a benchmark, the trees generated a factor that served as a hedge against the market. The resulting portfolios require less turnover and achieve higher Sharpe ratios compared to the original trees. This demonstrated the large potential of the trees to augment any existing model by forming a macroeconomic factor that captures macroeconomic signals and forces that were not yet captured by the existing model.

Additionally, the concept of random forests was applied to the macroeconomic trees, generating a large number of trees using bootstrapped data and a random susbet of macroeconomic features. This approach aimed to assess whether bagging could improve the out-of-sample performance of the tree portfolios. The macroeconomic random forests exhibited lower turnover due to averaging multiple portfolios into one, leading to slightly improved out-of-sample Sharpe ratios. However, transaction costs remained a challenge in maintaining Sharpe ratios in the out-of-sample period. The random forests also facilitated the quantification of macroeconomic feature importance, with the term spread and market return identified as the most crucial variables in the model.

Previous literature has highlighted the existence of a relation between recessions and expansions and the returns of several risk factors, a connection that the tree portfolios also appeared to capture by considering the market as a significant variables and incorporating the term spread as a anticipatory indicator. However, some other macroeconomic features examined in this study, previously associated with significant relationships, did not exert a significant impact on the model (e.g., industrial production, GDP, and inflation).

One of the main limitations of this study is the high computational power required to grow the trees, particularly for the random forest in combination with subset resampling. Consequently, not all forests reached full growth according to the stopping criterion, and extensive parameter tuning for random forests was limited. Therefore, it is still possible that the out-of-sample performance of these models can be improved given more resources. Additionally, the results indicated that most tree portfolios required very high turnover. This implies the need for extensive trading to capture macroeconomic signals. Future research could try to create portfolios using decision trees and macroeconomic variables while explicitly limiting turnover by for example penalising turnover, or training the model on the transaction-cost adjusted Sharpe ratio. Additionally, investigating the inclusion of additional macroeconomic variables, such as the default spread and a recession indicator, might yield valuable signals to improve performance.

In summary, this study demonstrated the ability of flexible tree structures to capture the relationships between macroeconomic variables and asset returns, enabling the construction of portfolios with higher Sharpe ratios accordingly. In doing so, this paper has tried to shine a light on the application of decision trees to portfolio construction and the creation of economic regimes, resulting in high-turnover portfolios with enhanced Sharpe ratios. However, further research is required to address computational challenges, reduce turnover, and explore the inclusion of additional macroeconomic variables. By advancing these methodologies and broadening its application, investors can benefit from improved portfolio performance and a deeper understanding of the complex dynamics between macroeconomics and financial markets.

References

Amenc, N., Esakia, M., Goltz, F. & Luyten, B. (2019). Macroeconomic risks in equity factor investing. *The Journal of Portfolio Management*, *45*(6), 39–60. doi: 10.3905/jpm.2019.1.092

Ammann, M. & Verhofen, M. (2006). The effect of market regimes on style allocation. *Financial Markets and Portfolio Management*, *20*(3), 309–337. doi: 10.1007/s11408-006-0018-2

Athanassakos, G. (2009). Value versus growth stock returns and the value premium: The canadian experience 1985-2005. *Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration*, *26*(2), 109–121. doi: 10.1002/cjas.93

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. doi: 10.1016/0304-405x(81)90018-0

Basak, S., Kar, S., Saha, S., Khaidem, L. & Dey, S. R. (2019). Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*, *47*, 552–567. doi: 10.1016/j.najef.2018.06.013

Benartzi, S., Michaely, R. & Thaler, R. (1997). Do changes in dividends signal the future or the past? *The Journal of Finance*, *52*(3), 1007–1034. doi: 10.1111/j.1540-6261.1997.tb02723.x

Bender, J., Briand, R., Melas, D. & Subramanian, R. A. (2013). Foundations of factor investing. *MSCI Research Insight*.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. doi: 10.1007/bf00058655

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. doi: 10.1023/a:1010933404324

Breiman, L., Friedman, J., Stone, C. & Olshen, R. (1984). *Classification and regression trees*. Taylor & Francis. Retrieved from `https://books.google.nl/books?id=JwQx-WOmSyQC`

Bryzgalova, S., Pelger, M. & Zhu, J. (2021). Forest through the trees: Building cross-sections of stock returns. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3493458

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82. doi: 10.1111/j.1540-6261.1997.tb03808.x

Chang, K.-L. (2009). Do macroeconomic variables have regime-dependent effects on stock return dynamics? evidence from the markov regime switching model. *Economic Modelling*, *26*(6), 1283–1299. doi: 10.1016/j.econmod.2009.06.003

Chen, N.-F., Roll, R. & Ross, S. A. (1986). Economic forces and the stock market. *The Journal of Business*, *59*(3), 383. doi: 10.1086/296344

Cong, L. W., Feng, G., He, J. & He, X. (2021). Asset pricing with panel trees under global split criteria. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3949463

De Bondt, W. F. & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, *40*(3), 793–805. doi: 10.1111/j.1540-6261.1985.tb05004.x

DeMiguel, V., Garlappi, L. & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, *22*(5), 1915–1953. doi: 10.1093/rfs/hhm075

Estrella, A. & Mishkin, F. S. (1996). The yield curve as a predictor of u.s. recessions. *Current Issues in Economics and Finance*, *2*(7). doi: 10.2139/ssrn.1001228

Fama, E. F. & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427–465. doi: 10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F. & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. doi: 10.1016/j.jfineco.2014.10.010

Foglia, M., Recchioni, M. C. & Polinesi, G. (2021). Smart beta allocation and macroeconomic variables: The impact of covid-19. *Risks*, *9*(2), 34. doi: 10.3390/risks9020034

Geurts, P., Irrthum, A. & Wehenkel, L. (2009). Supervised learning with decision tree-based methods in computational and systems biology. *Molecular BioSystems*, *5*(12), 1593. doi: 10.1039/b907946g

Gkatzilakis, G.-X. & Sivasubramanian, S. (2014). Active allocation of smart beta indices based on factor timing and regime switching. *EDHEC Business School*.

Hamilton, J. D. (2005). Regime switching models. *The New Palgrave Dictionary of Economics*, 1–7. doi: 10.1057/978-1-349-95121-$5_2459 - 1$

Hammerschmid, R. & Lohre, H. (2018). Regime shifts and stock return predictability. *International Review of Economics amp; Finance*, *56*, 138–160. doi: 10.1016/j.iref.2017.10.021

Harvey, C. R., Liu, Y. & Zhu, H. (2015). . . . and the cross-section of expected returns. *Review of Financial Studies*, *29*(1), 5–68. doi: 10.1093/rfs/hhv059

Humpe, A. & Macmillan, P. (2007). Can macroeconomic variables explain long term stock market movements? a comparison of the us and japan. *SSRN Electronic Journal*. doi: 10.2139/ssrn.1026219

Jareño, F. & Negrut, L. (2015). Us stock market and macroeconomic factors. *Journal of Applied Business Research (JABR)*, *32*(1), 325. doi: 10.19030/jabr.v32i1.9541

Jegadeesh, N. & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91. doi: 10.1111/j.1540-6261.1993.tb04702.x

Kan, R. & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, *42*(3), 621–656. doi: 10.1017/s0022109000004129

Kao, D.-L. & Shumaker, R. D. (1999). Equity style timing. *Financial Analysts Journal*, *55*(1), 37–48. Retrieved 2023-05-22, from `http://www.jstor.org/stable/4480137`

Khaidem, L., Saha, S. & Dey, S. R. (2016). Predicting the direction of stock market prices using random forest. *CoRR*, *abs/1605.00003*. Retrieved from `http://arxiv.org/abs/1605.00003`

Kim, M. K. & Burnie, D. A. (2002). The firm size effect and the economic cycle. *Journal of Financial Research*, *25*(1), 111–124. doi: 10.1111/1475-6803.00007

Kuosmanen, P. & Vataja, J. (2011). The role of stock markets vs. the term spread in forecasting macrovariables in finland. *The Quarterly Review of Economics and Finance*, *51*(2), 124-132. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1062976911000068` doi: https://doi.org/10.1016/j.qref.2011.01.005

Ledoit, O. & Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio.

Journal of Empirical Finance, *15*(5), 850–859. doi: 10.1016/j.jempfin.2008.03.002

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77. doi: 10.2307/2975974

Pastor, L. & Stambaugh, R. (2001). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*(3). doi: 10.3386/w8462

Prombutr, W., Phengpis, C. & Zhang, Y. (2012). What explains the investment growth anomaly? *Journal of Banking amp; Finance*, *36*(9), 2532–2542. doi: 10.1016/j.jbankfin.2012.05.010

Saha, A., Basu, S. & Datta, A. (2021). Random forests for spatially dependent data. *Journal of the American Statistical Association*, *118*(541), 665–683. doi: 10.1080/01621459.2021.1950003

Scherer, B. (2002). Portfolio resampling: Review and critique. *Financial Analysts Journal*, *58*(6), 98–109. doi: 10.2469/faj.v58.n6.2489

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The Journal of Finance*, *19*(3), 425–442. doi: 10.1111/j.1540-6261.1964.tb02865.x

Shen, W. & Wang, J. (2017). Portfolio selection via subset resampling. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*(1). doi: 10.1609/aaai.v31i1.10728

Singh, G. (2022). Machine learning models in stock market prediction. *International Journal of Innovative Technology and Exploring Engineering*, *11*(3), 2278–3075.

Tangjitprom, N. (2012). The review of macroeconomic factors and stock returns. *International Business Research*, *5*(8). doi: 10.5539/ibr.v5n8p107

Titman, S., Wei, K. J. & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, *39*(4), 677–700. doi: https://doi.org/10.1017/S0022109000003173

Ung, D. & Luk, P. (2016). What is in your smart beta portfolio? a fundamental and macroeconomic analysis. *Samp;P Dow Jones Indices*.

A Tables

Table A1: Summary statistics of the macroeconomic variables used to create trees

Abbreviation	Description	Mean	St Dev	Min	Max	Num. Obs
INF	Inflation	0.320	0.320	-1.771	1.810	718
GNP	GNP growth rate	3.000	4.437	-29.9	35.3	239
UNEM	Change in unemployment rate	0.140	9.345	-17.65	234.1	718
IND	Change in industrial production	0.191	0.950	-13.38	6.501	718
M1	Change in M1 money stock	0.841	8.939	-3.318	238.8	718
TS	Term spread	1.033	1.635	-6.51	3.85	718
PS	Pastor-Stambaugh illiqudity	-0.027	0.062	-0.462	0.198	718
MKT	S&P500 return	0.628	3.557	-20.39	12.02	718
RFR	Risk free rate	0.362	0.267	0	1.35	718

Table A2: Correlation matrix for the macroeconomic variables.

	MKT	INF	IND	M1	UNEM	GDP	RFR	PS	TS
MKT	1								
INF	-0.177	1							
IND	-0.033	-0.095	1						
M1	0.034	0.028	-0.039	1					
UNEM	-0.084	0.115	-0.416	-0.050	1				
GDP	0.046	-0.177	0.474	0.023	-0.402	1			
RFR	-0.077	0.517	-0.179	-0.031	0.148	-0.212	1		
PS	0.350	-0.139	0.024	-0.008	-0.146	0.120	-0.079	1	
TS	0.132	-0.488	0.176	0.195	-0.180	0.239	-0.604	0.201	1

Table A3: Summary statistics for all the investable assets, including the Jarque-Bera test statistic for normality.

Number	Factor	Portfolio	Mean	StDv	Min	Max	JB
1	**Size**	**Bottom 30%**	0.976	6.202	-29.49	26.67	152.3
2		**Medium 40%**	0.889	5.402	-27.30	22.45	177.4
3		**Top 30%**	0.662	3.330	-21.03	17.44	93.0
4	**Value**	**Bottom 30%**	0.917	4.706	-24.28	20.97	102.5
5		**Medium 40%**	0.941	4.368	-21.22	17.96	214.3
6		**Top 30%**	1.164	5.069	-27.38	24.01	411.7
7	**Profitability**	**Bottom 30%**	0.771	5.274	-25.72	16.79	150.7
8		**Medium 40%**	0.906	4.400	-20.94	17.43	1187.9
9		**Top 30%**	1.028	4.433	-22.63	16.68	106.6
10	**Investment**	**Bottom 30%**	1.106	4.449	-22.22	16.80	114.2
11		**Medium 40%**	0.941	4.141	-20.49	15.22	167.6
12		**Top 30%**	0.888	5.233	-25.02	20.94	95.2
13	**Earnings-to-Price**	**Bottom 30%**	0.883	4.816	-24.26	21.40	93.8
14		**Medium 40%**	0.956	4.249	-22.02	16.79	136.2
15		**Top 30%**	1.156	4.765	-24.09	25.13	309.5
16		**Negative**	1.033	7.352	-30.58	36.70	151.3
17	**Cashflow-to-Price**	**Bottom 30%**	0.916	4.823	-23.70	21.46	82.3
18		**Medium 40%**	0.929	4.329	-23.92	16.22	179.3
19		**Top 30%**	1.117	4.601	-24.09	24.08	345.9
20		**Negative**	0.958	7.527	-29.98	40.32	191.5
21	**Dividend Yield**	**Bottom 30%**	0.941	5.082	-25.36	21.46	121.2
22		**Medium 40%**	0.948	4.304	-23.33	15.29	145.0
23		**Top 30%**	1.00	4.043	-16.81	21.85	259.8
24		**Negative**	1.013	6.757	-30.48	27.42	73.6
25	**Short-Term Reversal**	**Bottom 10%**	1.014	7.651	-34.49	51.17	777.7
26		**Top 10%**	0.720	5.532	-27.12	24.45	112.1
27	**Momentum**	**Bottom 10%**	0.270	8.390	-27.98	44.89	644.3
28		**Top 10%**	1.451	6.119	-26.78	21.79	101.7
29	**Long-Term Reversal**	**Bottom 10%**	1.245	6.871	-29.97	39.00	252.0
30		**Top 10%**	0.924	5.887	-25.07	25.11	66.1
31	**Accruals**	**Bottom 20%**	1.112	5.135	-23.74	18.50	92.8
32		**Top 20%**	0.818	5.377	-29.05	21.94	139.3
33	**Market Beta**	**Bottom 20%**	0.888	3.520	-13.72	18.59	130.3
34		**Top 20%**	1.028	7.264	-31.01	32.57	60.7
35	**Net Share Issues**	**Bottom 20%**	0.918	4.303	-20.61	17.66	149.9
36		**Top 20%**	0.601	5.486	-25.22	19.07	133.3
37		**Negative**	1.102	4.260	-20.66	15.88	133.0
38		**Zero**	1.097	4.372	-19.74	16.92	129.4
39	**Variance**	**Bottom 20%**	0.945	3.508	-16.34	13.11	92.8
40		**Top 20%**	0.689	8.039	-30.90	28.59	59.4
41	**Residual Variance FF3**	**Bottom 20%**	0.964	3.794	-17.88	14.82	94.6
42		**Top 20%**	0.680	7.908	-30.83	35.60	108.8

B Figures

Table A4: Summary statistics for the average weights produced by the fully grown trees. Including the macroeconomic feature each variable has the highest correlation with.

	Portfolio Description	Mean	Max	St. Dev	Highest Correlation
1	**Small Size**	0.015	0.228	0.036	TS; 0.246
2	**Medium Size**	0.039	0.129	0.018	TS; 0.291
3	**Big Size**	0.014	0.058	0.007	MKT; 0.168
4	**Low Value**	0.003	0.082	0.010	TS; -0.178
5	**Medium Value**	0.029	0.242	0.027	TS; 0.201
6	**High Value**	0.091	0.241	0.063	TS; -0.305
7	**Low Profitability**	0.024	0.059	0.005	INF; -0.083
8	**Medium Profitability**	0.009	0.054	0.010	RFR; -0.065
9	**High Profitability**	0.011	0.129	0.019	TS; -0.094
10	**Low Investment**	0.052	0.197	0.028	TS; -0.402
11	**Medium Investment**	0.018	0.126	0.018	TS; -0.102
12	**High Investment**	0.046	0.111	0.009	INF; -0.171
13	**Low E/P**	0.002	0.065	0.007	TS; -0.173
14	**Medium E/P**	0.011	0.092	0.013	TS; -0.205
15	**High E/P**	0.085	0.234	0.058	TS; 0.342
16	**Negative E/P**	0.038	0.098	0.007	GDP; 0.148
17	**Low CF/P**	0.003	0.048	0.008	TS; -0.147
18	**Medium CF/P**	0.048	0.199	0.017	MKT; -0.116
19	**High CF/P**	0.087	0.217	0.049	TS; -0.157
20	**Negative CF/P**	0.024	0.059	0.004	INF; -0.104
21	**Low Dividend**	0.029	0.101	0.011	MKT; 0.106
22	**Medium Dividend**	0.049	0.137	0.016	TS; 0.150
23	**High Dividend**	0.111	0.231	0.061	MKT; -0.204
24	**Negative Dividend**	0.029	0.081	0.008	TS; 0.257
25	**Low ST Reversal**	0.045	0.187	0.015	MKT; 0.083
26	**High ST Reversal**	0.027	0.032	0.004	TS; -0.107
27	**Low Momentum**	0.032	0.034	0.004	TS; 0.221
28	**High Momentum**	0.067	0.206	0.050	MKT; 0.204
29	**Low LT Reversal**	0.048	0.213	0.023	MKT; 0.268
30	**High LT Reversal**	0.049	0.080	0.007	TS; 0.250
31	**Low Accruals**	0.027	0.166	0.020	MKT; 0.117
32	**High Accruals**	0.045	0.049	0.007	TS; 0.107
33	**Low Beta**	0.091	0.275	0.052	MKT; -0.129
34	**High Beta**	0.050	0.137	0.012	RFR; -0.097
35	**Low Net Share Issues**	0.052	0.211	0.030	RFR; -0.201
36	**High Net Share Issues**	0.021	0.038	0.004	TS; 0.198
37	**Negative Net Share Issues**	0.107	0.257	0.060	TS; -0.163
38	**Zero Net Share Issues**	0.048	0.230	0.077	TS; 0.439
39	**Low Variance**	0.064	0.205	0.044	TS; 0.262
40	**High Variance**	0.047	0.112	0.010	TS; 0.254
41	**Low Residual Variance**	0.057	0.231	0.044	TS; -0.207
42	**High Residual Variance**	0.046	0.146	0.015	TS; 0.259

Table A5: Performance measures for the portfolios obtained with market-boosted macroeconomic trees for a mean-variance portfolio of the market and the tree factor, in-sample as well as out-of-sample.

Panel A: In-sample (1963-2005)

Market-Boosted Macroeconomic Trees						
Cutpoints	**Splits**	**Avg Return**	**SR**	**Turnover**	α	**Net-SR**
2	**10**	1.261	1.472***	0.563	1.034***	1.213
	15	1.092	1.314***	0.210	0.846***	1.188
	20	1.301	1.563***	0.418	1.093***	1.311**
3	**10**	1.146	1.359***	0.313	0.904***	1.175
	15	1.149	1.362***	0.308	1.180***	1.180
	20	1.119	1.328***	0.206	0.872***	1.209*
4	**10**	1.287	1.591***	0.597	1.089***	1.220
	15	1.316	1.661***	0.784	1.130***	1.170
	20	1.464	1.887***	0.924	1.304***	1.298**
Benchmark Models						
Equally-weighted		0.971	0.695	0.041	0.942***	0.681
Value-weighted		0.910	0.711	0.055	0.892***	0.690
Mean-variance		1.272	1.080	0.036	1.255***	1.064

Panel B: Out-of-sample (2005-2023)

Market-Boosted Macroeconomic Trees						
Cutpoints	**Splits**	**Avg Return**	**SR**	**Turnover**	α	**Net-SR**
2	**10**	0.732	0.770	0.511	0.544**	0.499
	15	0.882	0.966**	0.309	0.670**	0.797
	20	0.755	0.781	0.577	0.594**	0.480
3	**10**	0.836	0.863	0.376	0.615**	0.666
	15	0.847	0.864	0.384	0.627**	0.666
	20	0.841	0.896*	0.261	0.625**	0.755
4	**10**	0.972	1.002***	0.543	0.765***	0.722
	15	0.824	0.880**	0.811	0.688**	0.446
	20	0.856	0.870*	0.915	0.810***	0.404
Benchmark Models						
Equally-weighted		0.871	0.601	0.041	0.873***	0.587
Value-weighted		0.865	0.673	0.056	0.875***	0.652
Mean-variance		0.836	0.660	0.037	0.856***	0.646

Note. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Net-SR is defined as the Sharpe Ratio net of transaction costs (based on the return series in Equation 12), under the assumption that holding the market is costless. The Sharpe ratios are compared against the mean-variance portfolios using the Ledoit-Wolf Sharpe ratio test. All trees were grown with $s = 500$ and $b = 10$ for subset resampling.

Table A6: Performance measures for the portfolios obtained with market-boosted macroeconomic trees for a 50-50 portfolio of the market and the tree factor, in-sample as well as out-of-sample.

Panel A: In-sample (1963-2005)

Market-Boosted Macroeconomic Trees						
Cutpoints	**Splits**	**Avg Return**	**SR**	**Turnover**	α	**Net-SR**
2	**10**	1.090	1.396**	0.411	0.755***	1.180
	15	0.994	1.275	0.165	0.667***	1.170
	20	1.092	1.447**	0.311	0.755***	1.244
3	**10**	1.037	1.320**	0.247	0.714***	1.164
	15	1.039	1.322**	0.243	0.716***	1.169
	20	1.017	1.291*	0.163	0.692***	1.191
4	**10**	1.078	1.460**	0.408	0.743***	1.185
	15	1.105	1.526***	0.544	0.784***	1.155
	20	1.163	1.650***	0.592	0.835***	1.240
Benchmark Models						
Equally-weighted		0.971	0.695	0.041	0.942***	0.681
Value-weighted		0.910	0.711	0.055	0.892***	0.690
Mean-variance		1.272	1.080	0.036	1.255***	1.064

Panel B: Out-of-sample (2005-2023)

Market-Boosted Macroeconomic Trees						
Cutpoints	**Splits**	**Avg Return**	**SR**	**Turnover**	α	**Net-SR**
2	**10**	0.707	0.838	0.373	0.397**	0.615
	15	0.832	0.984*	0.244	0.520***	0.840
	20	0.719	0.866	0.399	0.410**	0.621
3	**10**	0.796	0.901	0.297	0.485**	0.728
	15	0.804	0.903	0.303	0.495**	0.729
	20	0.800	0.932	0.207	0.496**	0.805
4	**10**	0.866	1.041**	0.371	0.553***	0.817
	15	0.767	0.959*	0.562	0.477**	0.603
	20	0.777	0.981*	0.586	0.490***	0.607
Benchmark Models						
Equally-weighted		0.871	0.601	0.041	0.873***	0.587
Value-weighted		0.865	0.673	0.056	0.875***	0.652
Mean-variance		0.836	0.660	0.037	0.856***	0.646

Note. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Net-SR is defined as the Sharpe Ratio net of transaction costs (based on the return series in Equation 12), under the assumption that holding the market is costless. The Sharpe ratios are compared against the mean-variance portfolio in terms of the Sharpe ratio using the Ledoit-Wolf Sharpe ratio test. All trees were grown with $s = 500$ and $b = 10$ for subset resampling.

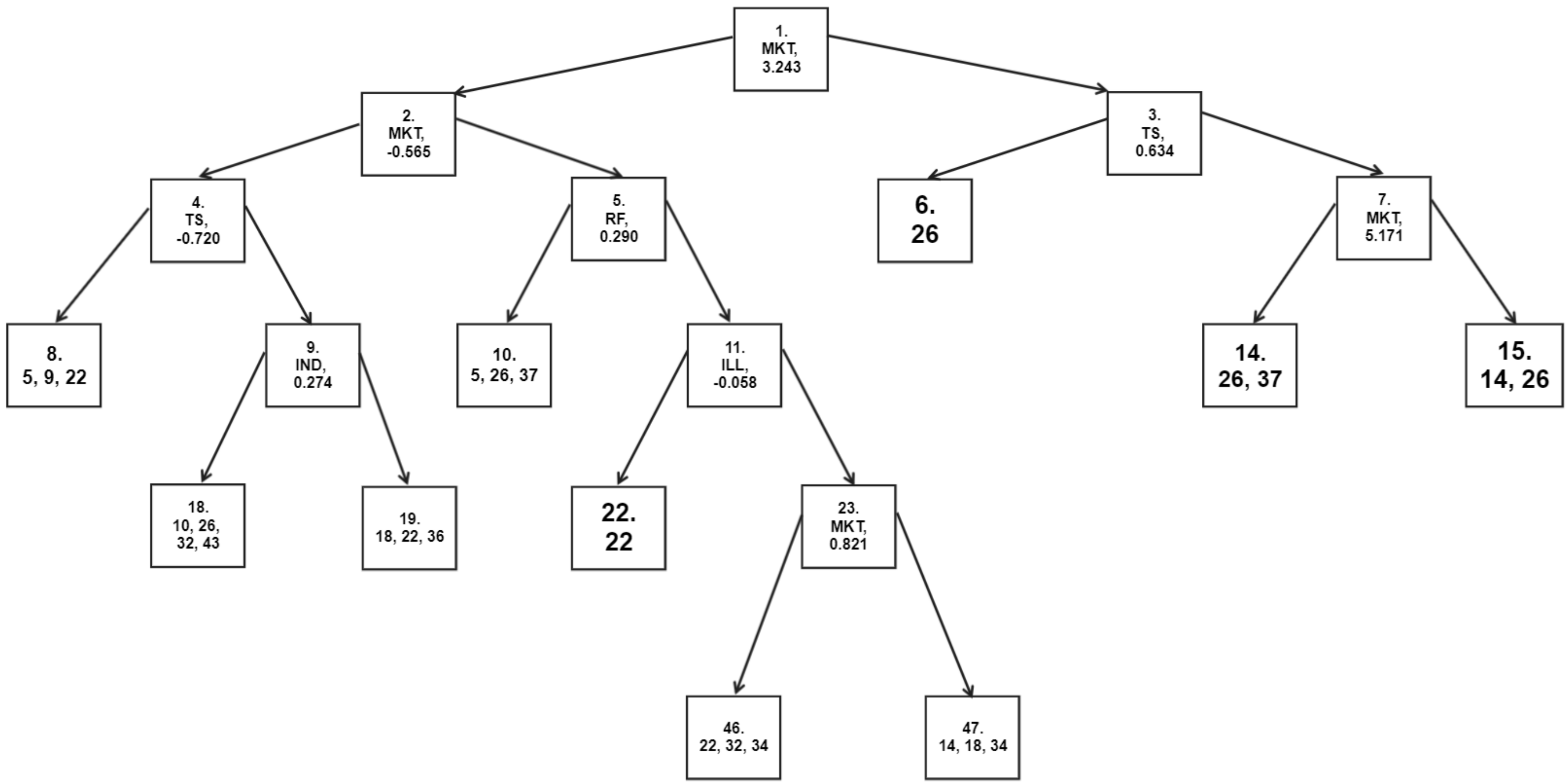


Figure B3: Fully grown tree using four cutpoints and at most fifteen splits. Each square represents a node containing the following information in case it is not a leaf node: the number of the node, the variable that was used to split on, and the split point. If the square represents a leaf node it contains the following information: the number of the node and the portfolios that receive a weight larger than 10%, where the numbers refer to the portfolios as in Table A3. The abbreviations of the macroeconomic variables are as follows: MKT, market, TS, term spread, RF, risk-free rate, IND, industrial production (see Table A1). All observations left of a certain node have values lower than the split point for the split variables, and all observations to the right have a higher value.

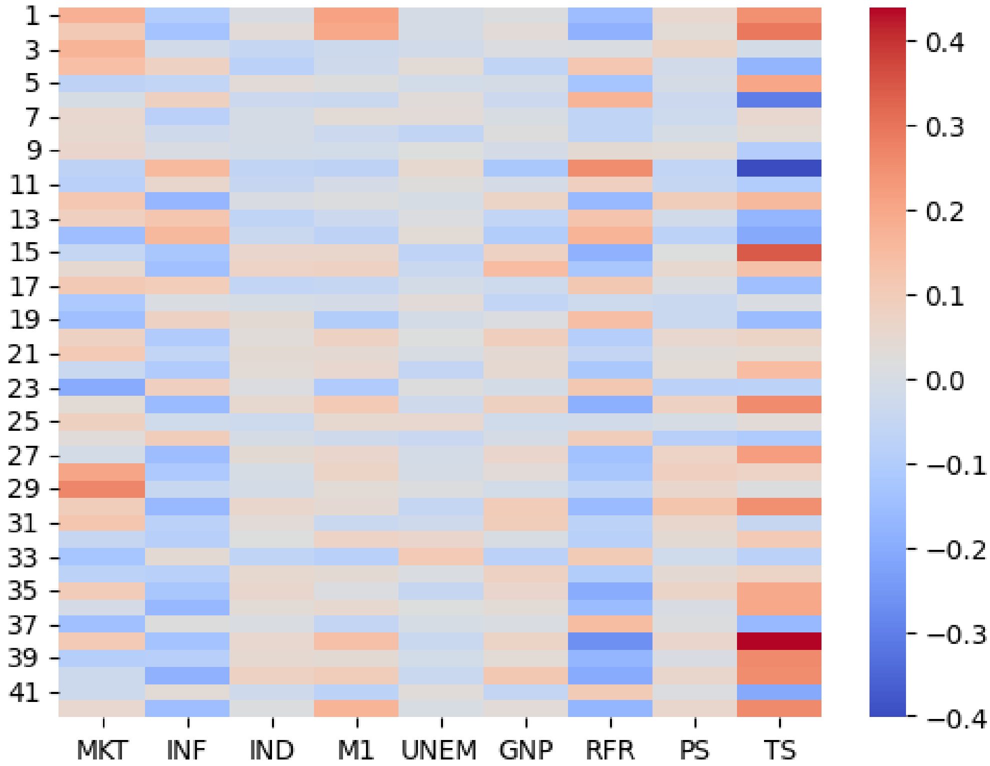
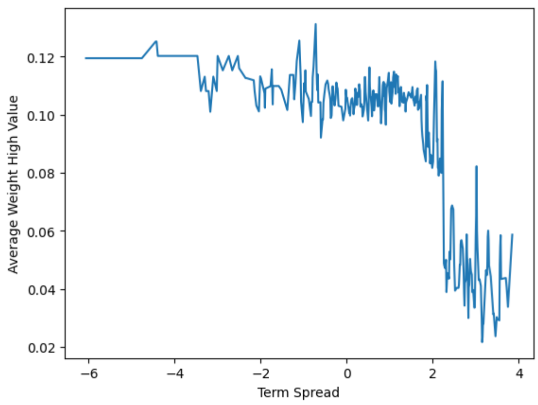
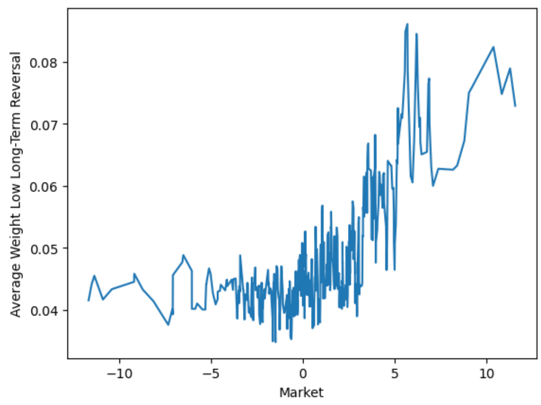


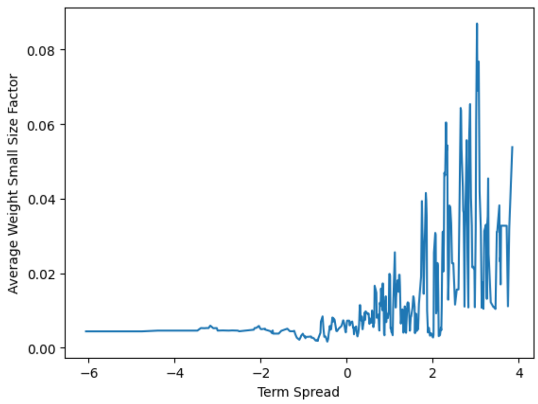
Figure B4: Correlations of the average weights for the factors and the macroeconomic variables for the nine fully grown trees. The abbreviations of the macroeconomic variables are as in Table A1, and the numbers of the portfolios are as in Table A3.



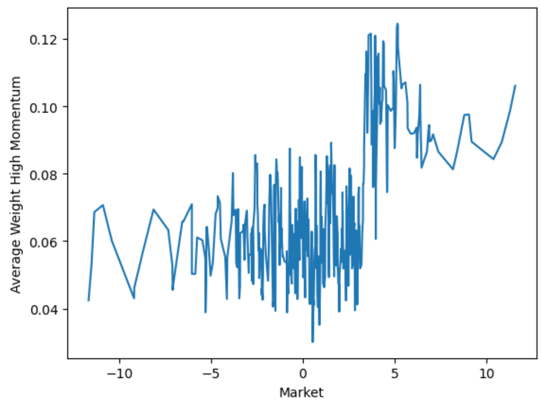
(a) Average portfolio weight of the high value factor as a function of the term spread.



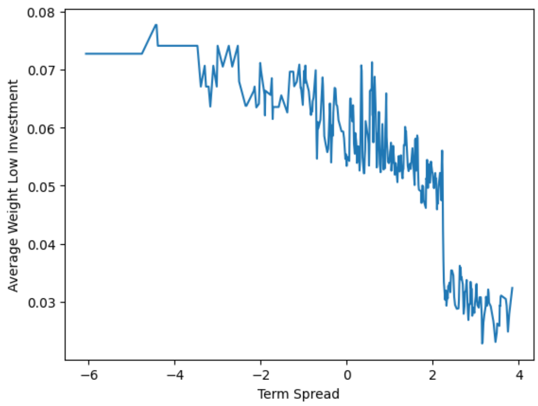
(b) Average portfolio weight of the low long-term reversal factor as a function of the market return.



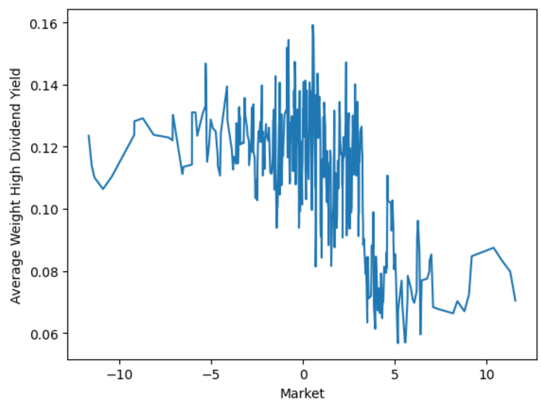
(c) Average portfolio weight of the small size factor as a function of the term spread.



(d) Average portfolio weight of the high momentum factor as a function of the market return.



(e) Average portfolio weight of the low investment factor as a function of the term spread.



(f) Average portfolio weight of the high dividend yield factor as a function of the market return.

Figure B5: The average portfolio weight of some factors as a function of the term spread and the market return.

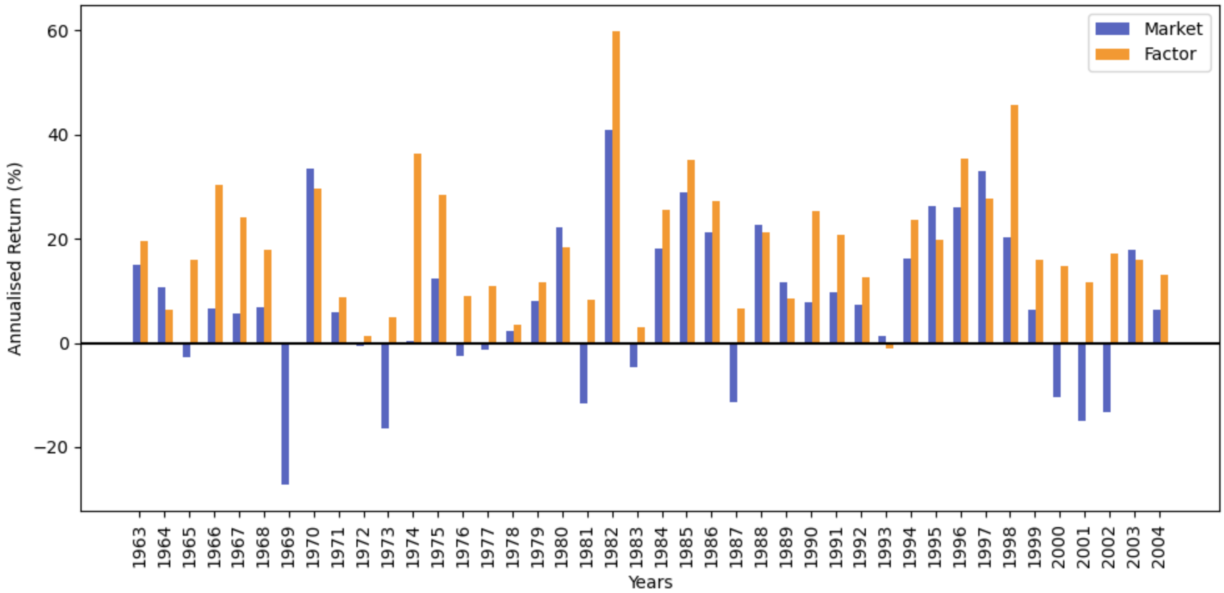
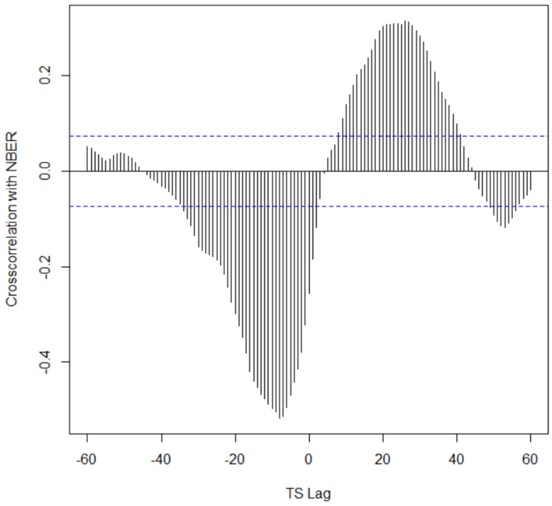
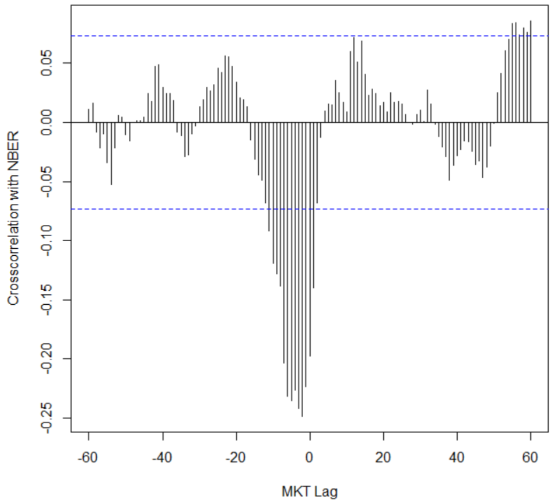


Figure B6: In-sample annualised returns for the market and the factor produced by the market-boosted tree for a tree using 4 cutpoints and 20 leafs.

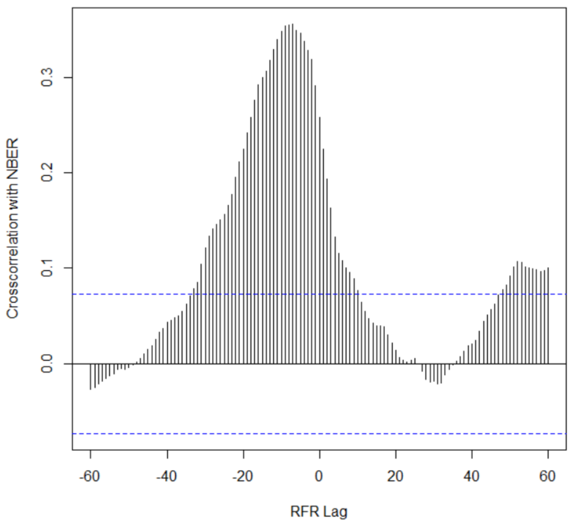
Figure B7: Out-of-sample annualised returns for the market and the factor produced by the market-boosted tree for a tree using 4 cutpoints and 20 leafs.



(a) Term Spread



(b) Market



(c) Risk-free-rate

Figure B8: Monthly crosscorrelation functions for three of the macroeconomic features with the NBER recession indicator.