TU Dortmund

Case Studies

Project III: Forecasting The Equity Premium Comparisons And Conclusions

Lecturers:
Prof. Dr. Matei Demetrescu
Dr. Paul Navas

Author: Bushra Tariq Kiyani

Group number: 4

Group members: Sarmistha Bhattacharyya, Oliver Fischer

January 25, 2024

Contents

1 Introduction

In the dynamic realm of financial markets, the perpetual challenge for investors, analysts, and portfolio managers lies in constructing sophisticated predictive models. Central to this pursuit is the accurate forecasting of stock returns—a pivotal undertaking with the potential to redefine investment strategies and fortify risk management practices. Within this context, the estimation of the equity premium assumes paramount importance. This premium, delineating the excess return anticipated from investing in the stock market over the risk-free rate, underscores the criticality of precise predictions in optimizing asset allocation decisions and elevating overall portfolio performance. (Julio et al., 2022)

In the final stage of our research endeavor, we embark on an exploration that builds upon the foundations laid in our two previous projects. These projects, focusing on the forecasting of excess returns, took distinct approaches: the first utilizing linear methods and the second integrating machine learning techniques. As we approach the culmination of our study, our objective is to extend these outcomes by incorporating rolling-windows estimation, enabling a comprehensive comparison of all resulting forecasts. The ultimate aim is to draw overarching conclusions regarding the forecastability of excess returns while providing insightful explanations for the observed trends and patterns.

The methodology for this phase encompasses two primary dimensions. Firstly, we employ a lagged-predictor-only approach, generating forecasts based solely on historical values of excess returns. The second dimension involves an expansion of our predictive framework to include covariates, providing an enriched context for forecasting. In both instances, the models are updated in real-time, allowing us to assess their performance dynamically. To ensure consistency in our analysis, we have set a maximal lag of 10.

A critical aspect of our analysis lies in the evaluation of forecast quality. We utilize the root mean squared forecast error (RMSFE), a metric that has guided our assessment in previous stages. However, in this phase, our evaluation goes beyond conventional metrics, incorporating the Diebold-Mariano Test. We will apply the Diebold-Mariano Test to each pair of forecasts derived from the initial modeling exercises, discerning the models that demonstrate superior performance. These findings, when amalgamated with insights from prior projects and scholarly literature, will pave the way for conclusive remarks on the efficacy of our forecasting models, shedding light on their strengths, weaknesses, potential improvements, and providing a broader perspective for future research in this domain.

The second section describes the structure and quality of the dataset in more detail. Additionally, the goals of the project are stated in the second section. The third section explains the statistical methods. The fourth section focuses on the application of these methods and the interpretation of the plots and results. Finally, the fifth section summarizes the most important findings.

2 Problem statement

2.1 Project Objectives

In the upcoming phase of our research project, our primary objectives revolve around advancing our forecasting methodologies for excess returns. Our foremost goal is to extend the existing forecasting models by implementing a dynamic rolling-windows estimation approach. This strategic enhancement ensures that our models adapt over time, providing a more responsive depiction of evolving market dynamics. Simultaneously, we will focus on generating one-step-ahead forecasts using both lagged predictors and covariates. The models will be systematically reestimated at each time point, leveraging the latest available information to enhance the precision of our forecasts. By incorporating this approach, we anticipate gaining valuable insights into the relative performance of these models and their effectiveness in capturing the complexities of excess return dynamics.

In the subsequent phase, our key objective is to conduct a comprehensive evaluation of the forecasting models, incorporating the Root Mean Squared Forecast Error (RMSFE) as a crucial metric. The RMSFE, along with the Diebold-Mariano Test, will be employed to rigorously compare the forecast performances of different models, considering both the magnitude and statistical significance of differences in forecast accuracy. This combined evaluation strategy aims to identify models that excel in predicting excess returns. Through these comparative analyses, we seek to discern patterns, strengths, and potential weaknesses in the various forecasting approaches employed. Ultimately, our overarching goal is to draw meaningful conclusions about the forecastability of excess returns, providing actionable insights for refining investment strategies and fortifying risk management practices in the dynamic landscape of financial markets.

2.2 Dataset and Data Quality

This project relies on the dataset elucidated in the initial report. The dataset, sourced from Amit Goyal's webpage and updated until 2022 (Goyal, 2022), covers the period from 1871 to 2022, with a primary focus on the equity premium as the dependent variable. This premium is meticulously computed and further categorized into three distinct time intervals: monthly, quarterly, and yearly. Notably, the dataset exhibits variations in the number of independent variables and observations across these time spans. Specifically, for monthly data, there are 18 independent variables with 1824 observations; for quarterly data, 22 covariates are considered alongside 608 observations; and for yearly data, 21 covariates are included with 152 observations. A detailed description of all variables is presented in Table 4 of the first report's appendix.

This dataset, curated for scientific investigation, is anticipated to meet high standards in terms of accuracy, consistency, relevance, and reliability as it is obtained from a reputable source. It is imperative to note that the presence of missing values in nearly all variables has been addressed by excluding them from the analysis, ensuring the integrity and reliability of the subsequent research findings.

3 Statistical methods

3.1 One Step Ahead Forecasts with Rolling Windows

The one-step-ahead forecasting approach serves as a cornerstone in time series analysis, pivotal for predicting the future values of a variable based on historical data. This methodology is especially indispensable in the dynamic realm of financial markets, where anticipating future values is fundamental for making well-informed decisions.

The one-step-ahead forecasting process involves predicting the future value of the dependent variable (y_t), at time $t+1$ based on the information available up to time t. The general equation for generating one-step-ahead forecasts is given by (Inoue et al., 2016):

$$\hat{y}_{t+1} = \hat{f}(y_t, \ldots, y_{t-n+1}, X_t)$$

where: $\hat{y}_{t+1}$ is the one-step-ahead forecast for the dependent variable at time $t+1$, $\hat{f}(\cdot)$ represents the estimated forecasting function, $y_t, \ldots, y_{t-n+1}$ are lagged values of the dependent variable, and X_t denotes relevant covariates and predictors. This dynamic forecasting approach ensures continuous model reestimation at each time point, incorporating the most recent available information for heightened precision in predicting future values.

Adaptation With Rolling Windows: In the one-step-ahead forecasting process, a rolling window strategy is employed to enhance forecast precision. Rolling windows involve reestimating the model at each time point, utilizing a fixed window of the most recent observations. This dynamic adaptation ensures responsiveness to short-term fluctuations in the data, providing a real-time adjustment to evolving market dynamics.

The general approach to one-step-ahead forecasting involves a systematic progression through key steps:

Model Specification: The process commences with a meticulous model specification, where appropriate statistical or machine learning techniques are selected based on the dataset's characteristics. This model captures historical relationships between the dependent variable and relevant predictors, aiming to discern patterns and dependencies crucial for future predictions.

Data Splitting: Following model specification, the dataset is strategically divided into two sets: the training set and the test set. The training set forms the basis for estimating the model parameters, allowing the model to learn underlying patterns and dependencies. The test set, unseen during model estimation, is reserved for evaluating the model's predictive performance.

Model Estimation: Next, the forecasting model is estimated using the training set, incorporating lagged values of the dependent variable and relevant covariates. This dynamic model is then applied to the test set to generate one-step-ahead forecasts, adapting in real-time to the most recent available information.

Evaluation Metrics: To assess the accuracy of one-step-ahead forecasts, evaluation metrics such as the Root Mean Squared Forecast Error (RMSFE), Mean Absolute Error (MAE), or Mean Squared Prediction Error (MSPE) are employed. These metrics quantify the disparities between predicted and observed values, offering a precise measure of forecast precision.

Iterative Process: Crucially, the one-step-ahead forecasting process is iterative. With the introduction of new observations, the model is continually updated and refined, enhancing its responsiveness and accuracy over time. This iterative nature makes it particularly well-suited for capturing the evolving dynamics inherent in financial markets and other time-dependent phenomena.

The one-step-ahead forecasting approach is a dynamic and iterative process that enables continuous refinement of predictions as new data becomes available. This methodology is essential in capturing the evolving nature of time series data, making it a valuable tool for decision-makers in various fields, particularly in finance and economics.

(Inoue et al., 2016)

3.2 The Diebold-Mariano Test

The Diebold-Mariano (DM) test is a statistical method designed to compare the forecasting accuracy of two competing models. Specifically, it assesses whether the difference in forecast errors between two models is statistically significant. The test is widely utilized in various fields, including finance and economics, where accurate predictions are paramount for informed decision-making. In forecasting studies, the DM test aids researchers in selecting the superior model among competing alternatives. It provides a robust statistical framework for evaluating whether the observed differences in forecast errors are beyond what would be expected by random chance.

Test Statistic: The Diebold-Mariano test statistic is defined as follows:

$$DM = \frac{\bar{\epsilon}_A - \bar{\epsilon}_B}{\sqrt{\left(\frac{\hat{\sigma}^2_{\epsilon,A} + \hat{\sigma}^2_{\epsilon,B} - 2\hat{\sigma}_{\epsilon,AB}}{T}\right)}}$$

where:

$\bar{\epsilon}_A$ and $\bar{\epsilon}_B$ are the mean forecast errors for models A and B, respectively.

$\hat{\sigma}^2_{\epsilon,A}$ and $\hat{\sigma}^2_{\epsilon,B}$ are the estimated variances of the forecast errors for models A and B, respectively.

$\hat{\sigma}_{\epsilon,AB}$ is the estimated covariance between the forecast errors of models A and B.

T is the number of observations.

Null Hypothesis (H_0): The null hypothesis posits that there is no difference in forecast accuracy between models A and B, i.e., $H_0 : \bar{\epsilon}_A = \bar{\epsilon}_B$.

Alternative Hypothesis (H_1): The alternative hypothesis suggests that there is a significant difference in forecast accuracy between models A and B, i.e., $H_1 : \bar{\epsilon}_A \neq \bar{\epsilon}_B$.

Critical Values: The test statistic is compared to critical values from the standard normal distribution to determine statistical significance. If the absolute value of the test statistic is greater than the critical value at a chosen significance level (e.g., 5%), the null hypothesis is rejected in favor of the alternative hypothesis.

Considerations: It's important to note that the DM test assumes that the forecast errors are uncorrelated and homoscedastic. Additionally, the test is sensitive to the choice of loss function used to compute forecast errors.

(Diebold and Mariano, 1995)

4 Statistical analysis

In this section, a comprehensive analysis of the dataset is presented using machine learning models discussed previously. For model training and testing, and graphical representations R software (R Core Team, 2022) Version, 4.2.1 is used with additional packages **dplyr** (Wickham et al., 2022), **ggplot2** (Wickham, 2016), **readr** (Wickham et al., 2023a), **rpart** (Therneau and Atkinson, 2022), **rpart.plot** (Milborrow, 2022), **randomForest** (Liaw and Wiener, 2022), and **caret** (Kuhn and Others, 2022), **ggpubr** (Kassambara, 2022), **tidyr** (Wickham et al., 2023b), **stats** (R Core Team, 2023), **forecast** (Hyndman et al., 2023), **pheatmap** (Kolde, 2019).

Three datasets are loaded from files named "PredictorData2022.xlsx - Monthly.csv," "PredictorData2022.xlsx - Quarterly.csv," and "PredictorData2022.xlsx - Annual.csv," respectively. Column names of each dataset are renamed for clarity and consistency and NA-values are removed from each dataset. The excess returns series is generated based on the stock returns and the risk-free rate. The excess returns are computed as the difference between the growth rates of the "Index" series and the corresponding lagged "Risk-free Rate" values. To conduct forecasts, the dataset is processed to include lagged excess returns and lagged predictors. The columns related to the stock returns, date, index, and risk-free rate are excluded as they are used to derive the dependent variable.

4.1 Dynamic One-Step-Ahead Forecasting Using Lagged Excess Returns: Incorporating Rolling Windows/Expanding Windows Approach

In this section, we employ a dynamic approach to generate one-step-ahead forecasts for excess returns based on lagged information. The forecasts are constructed using three distinct models: linear, regression tree, and random forest. The forecasting process is implemented at three different frequencies: annual, quarterly, and monthly. The objective is to assess the models' ability to capture evolving patterns and generate accurate predictions as new data becomes available. The forecasting models incorporate lagged excess returns as predictors.

We adopt a rolling window approach, updating the models at each forecasting point. The rolling window strategy allows for the dynamic adaptation of the models to the most recent available information, ensuring continuous refinement. To ensure a robust evaluation, we adhere to a split-sample approach. Specifically, we designate the last 10 years of data for out-of-sample forecasting and subsequent comparisons. This approach provides a consistent testing ground for all models, allowing us to assess their performance under similar market conditions.

The line charts presented below showcase the one-step-ahead forecasts of excess returns for each time within the designated test data generated by different models, providing a comparison against the actual observed data. The horizontal axis denotes various time points, while the vertical axis represents the values of excess returns. This analysis aims to assess the forecasting accuracy of each model against different temporal granularities: annual, quarterly, and monthly data.

Figure 1 illustrates the one-step-ahead forecasts by the Linear Model (in red), Regression Tree Model (in green), and Random Forest Model (in blue), juxtaposed with the observed excess returns (in black) across different time points. The Linear Model Forecasts and the Random Forest Forecasts exhibit relatively stable trends with slight fluctuations throughout the time points. Interestingly, the Regression Tree Forecasts mirrors the actual excess return but with less intensity. This suggests that while the Regression Tree Forecast is responsive to changes in the actual excess return, it may not fully capture the magnitude of these changes.

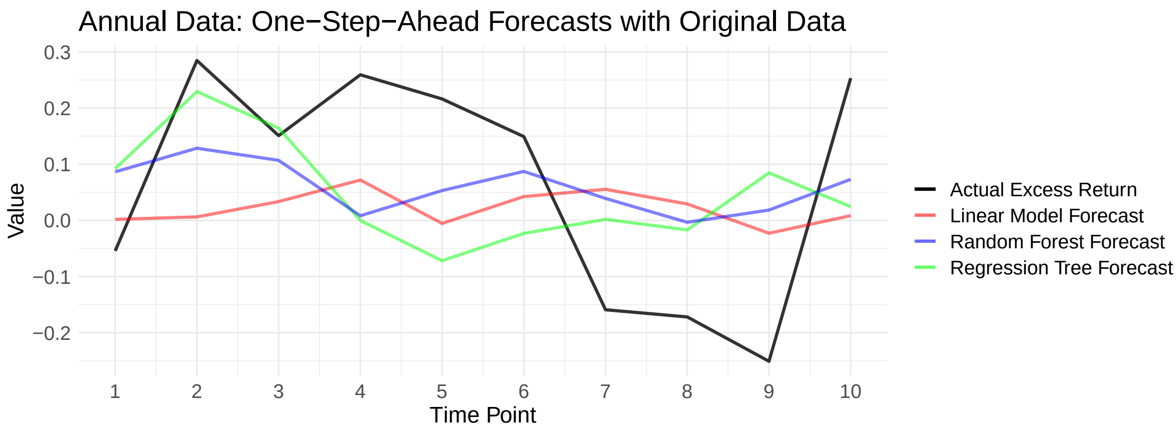

Figure 1: Comparison of One-Step-Ahead Excess Returns Forecasts – Annual Data

Figure 2 displays the one-step-ahead forecasts for quarterly data, providing a visual representation of the forecasting models' performance. The actual excess return is depicted by a black line, which fluctuates between positive and negative values across the time points. The Linear Model Forecast, represented by a red line, the Random Forest Forecast, represented by a blue line, and the Regression Tree Forecast, represented by a green line, all closely follow the trend of the actual excess return but with less fluctuation. This suggests that while these models are responsive to changes in the actual excess return, they may not fully capture the volatility of these changes.

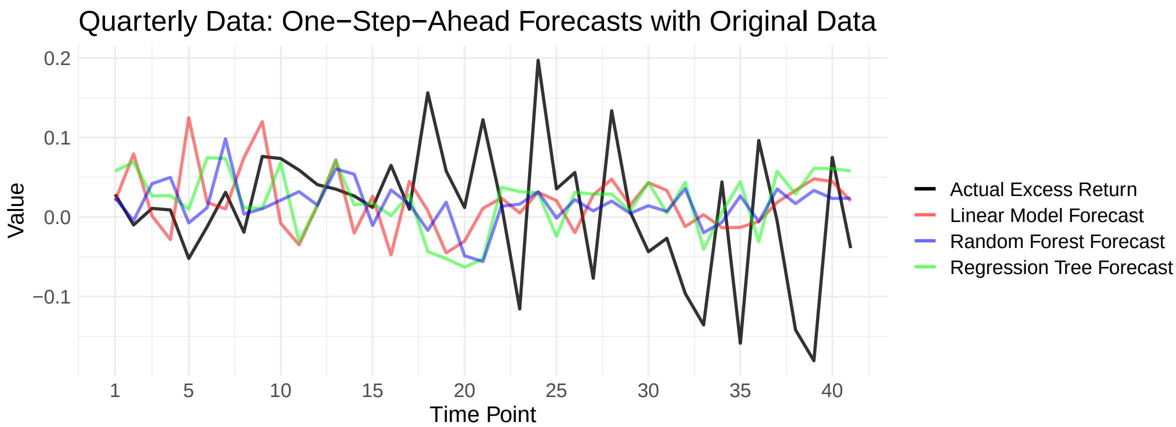

Figure 2: Comparison of One-Step-Ahead Excess Returns Forecasts – Quarterly Data

Figure 3 showcases the monthly data forecasts generated by the Linear Model, Regression Tree Model, and Random Forest Model, compared against the actual excess returns. The actual excess return is depicted by a black line, which fluctuates wildly across the entire range of time points. The Linear Model Forecast, represented by a red line, is a smoother line that follows a general trend but does not capture all fluctuations of the actual excess return. The Regression Tree Forecast, represented by a green line, is more erratic than the linear model but still smoother than the actual excess return, attempting to capture its fluctuations. The Random Forest Forecast, represented by a blue line, is similar in behavior to the regression tree but appears slightly smoother. This suggests that while these models are responsive to changes in the actual excess return, they may not fully capture the volatility of these changes.

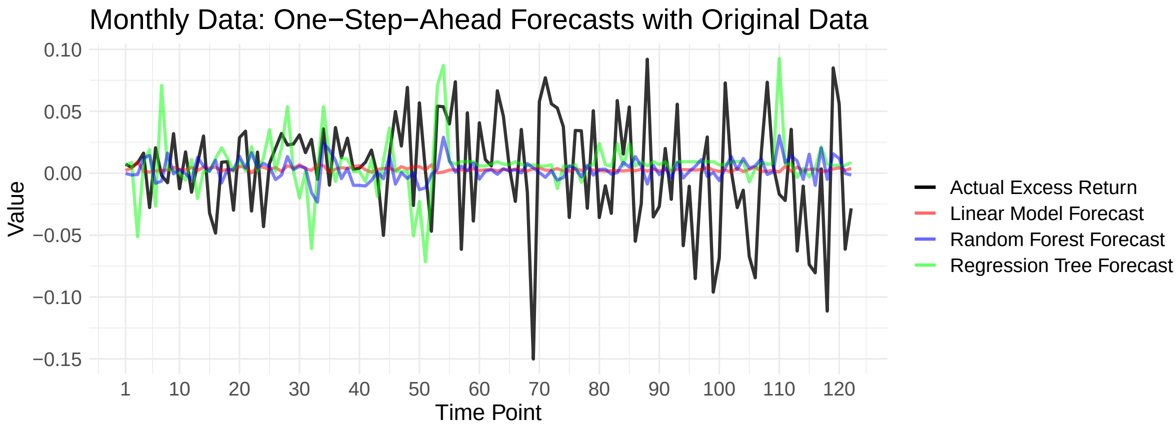

Figure 3: Comparison of One-Step-Ahead Excess Returns Forecasts – Monthly Data

Figure 4 presents the Root Mean Squared Forecast Error (RMSFE) values for Linear Model, Random Forest, and Regression Tree models fitted using lagged excess returns. In the annual data, the Linear Model exhibits the lowest RMSFE value (0.193), outperforming Random Forest (0.227) and Regression Tree (0.262). In the quarterly data, Regression Tree achieves the lowest RMSFE (0.087), followed by Random Forest (0.088) and Linear Model (0.091). Meanwhile, in the monthly data, the Linear Model again demonstrates the lowest RMSFE (0.044), surpassing Random Forest (0.045) and Regression Tree (0.052)

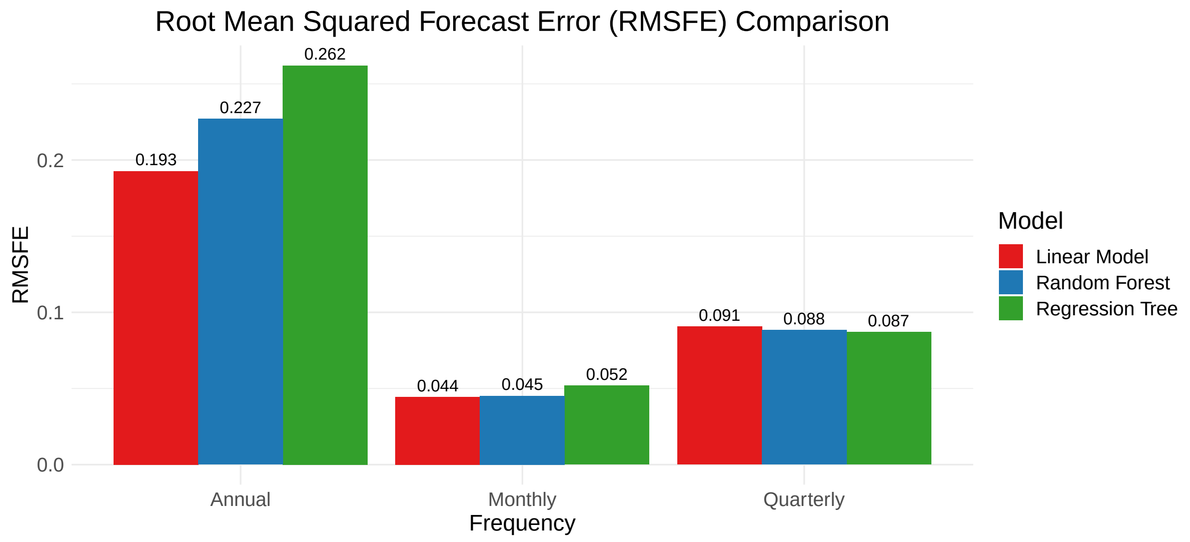

Figure 4: Comparison of Root Mean Squared Forecast Error (RMSFE) for Different Models Across Annual, Quarterly, and Monthly Frequencies: The bar chart illustrates the RMSFE values calculated for Linear Model, Regression Tree, and Random Forest predictions with the lagged excess return over the last 10 years.

4.2 Enhanced One-Step-Ahead Forecasting with Covariates: Incorporating Additional Predictors Alongside Excess Return Lags

The objective of this task is to refine the one-step-ahead forecasts for excess returns by incorporating additional covariates alongside lagged excess return variables. This process aims to enhance the forecasting accuracy of the models by considering a more comprehensive set of predictors. The following models are employed for this task: Linear Regression Model, Backward Selection Regression Model, Forward Selection Regression Model, Regression Tree Model and Random Forest Model.

For each of these models, forecasts are generated iteratively for annual, quarterly, and monthly datasets. The iterative process involves updating the model with additional observations at each step. The covariates included in each model differ; the linear model encompasses all available covariates, while the backward and forward selection models utilize the covariates selected in report 1 (refer to the table in the appendix for the selected covariates). In contrast, the decision tree and random forest models independently identify pertinent covariates based on their influence on excess returns.

Figure 5 shows a comparison of Root Mean Squared Forecast Errors (RMSFE) for various forecasting models for annual data. Notably, Linear Predictors model exhibits highest RMSFE of 0.3711, Tree Predictors model demonstrates second highest RMSFE of 0.2496, close to Backward Predictors model which exhibits RMSFE of 0.2463. Forward Predictors model performed better with an RMSFE of 0.2065. Random Forest Predictors model displays competitive performance with RMSFE value of 0.2322.

Additionally, among the individual lagged predictors, variables such as the Consumption Wealth Income Ratio (0.1976) and the lagged Inflation (0.1978) stand out for their relatively lower RMSFE values, indicating better predictive performance. These instances of lower RMSFE values highlight specific predictors and models that contribute to more precise and reliable forecasts in the given context. This comprehensive analysis provides a nuanced understanding of the forecasting models' effectiveness in capturing the dynamics of annual excess returns, shedding light on the key factors influencing their predictive accuracy.

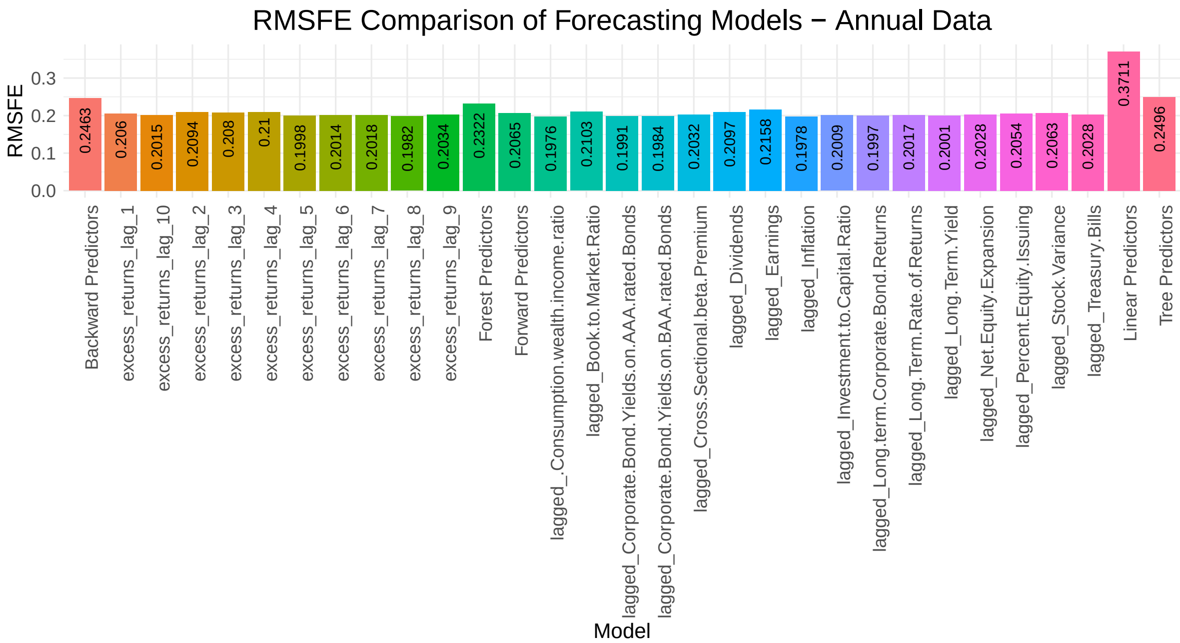

Figure 5: Comparison of Root Mean Squared Forecast Errors (RMSFE) for various forecasting models - Annual Data

Figure 6 shows the assessment of Root Mean Squared Forecast Errors (RMSFE) for quarterly excess return data, a closer examination reveals notable performance distinctions among various forecasting models. Linear Predictors, with an RMSFE of 0.3214, demonstrate a higher error compared to Backward Predictors (0.0974) and Forward Predictors (0.0882). Interestingly, both Tree Predictors (0.0853) and Forest Predictors (0.0889) exhibit competitive accuracy, with Tree Predictors achieving a particularly low RMSFE.

Examining the individual lagged predictors provides further insights. Notably, lagged Consumption Wealth Income Ratio stands out with an RMSFE of 0.0812, showcasing its effectiveness in contributing to more accurate quarterly forecasts. Other predictors, such as lagged Treasury Bills (0.0837) and lagged Net Equity Expansion (0.0841), also display relatively lower RMSFE values, emphasizing their positive impact on predictive performance.

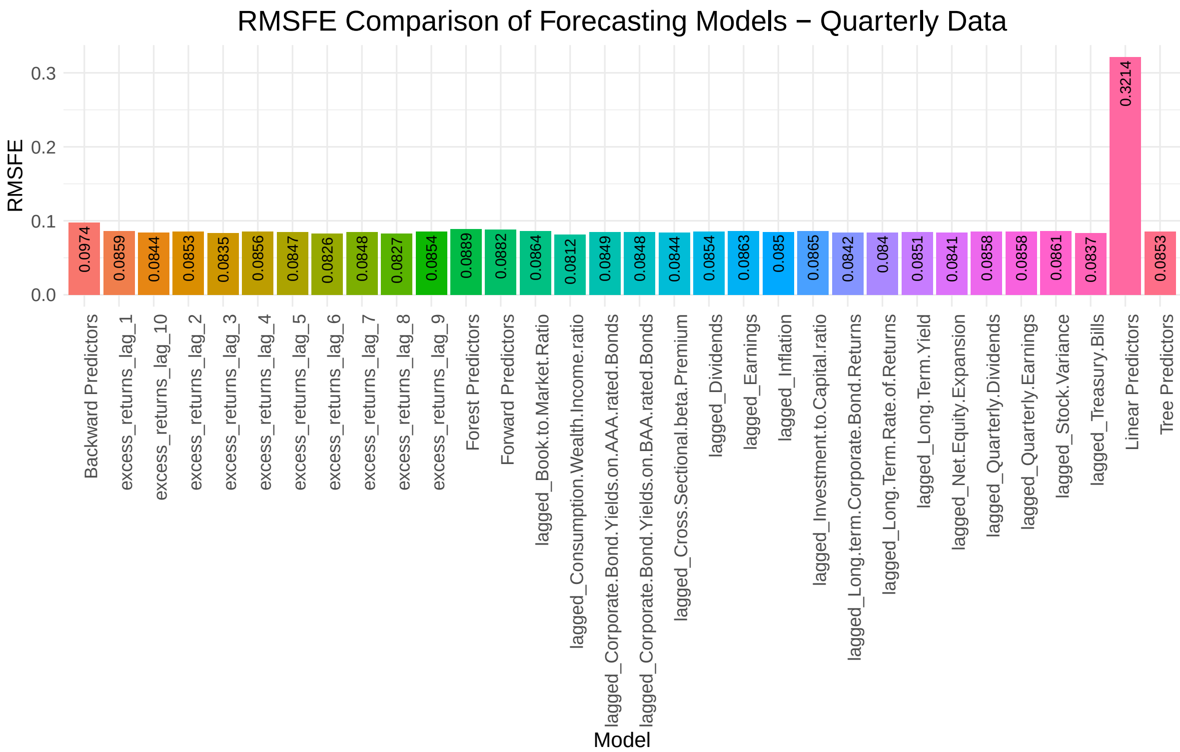

Figure 6: Comparison of Root Mean Squared Forecast Errors (RMSFE) for various forecasting models - Quarterly Data

Figure 7 shows models' RMSFEs comparison for monthly data. Notably, all predictor categories—Linear Predictors (0.0464), Backward Predictors (0.0457), and Forward Predictors (0.0457)—exhibit consistently low RMSFE values, indicating their strong accuracy in capturing the monthly excess return dynamics. Tree Predictors (0.0484) and Forest Predictors (0.0474) maintain competitive performance, with Tree Predictors slightly higher but still within an acceptable range.

Individual lagged predictors contribute significantly to the overall forecasting accuracy. Lagged predictors, such as lagged Inflation (0.0444), lagged Dividends (0.0446), and lagged Net Equity Expansion (0.0447), demonstrate particularly low RMSFE values, showcasing their effectiveness in enhancing the precision of monthly excess return forecasts.

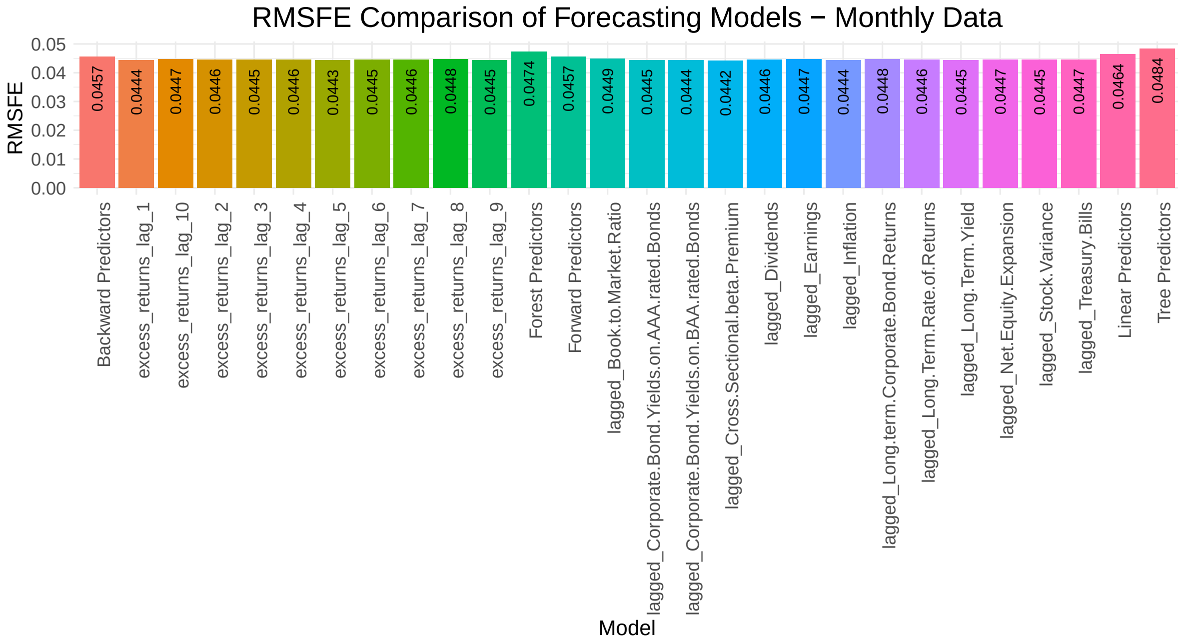

Figure 7: Comparison of Root Mean Squared Forecast Errors (RMSFE) for various forecasting models - Monthly Data

4.3 Diebold-Mariano Test for Model Comparison and Selection in Forecasting

In this project phase, the Diebold-Mariano Test is applied to systematically evaluate the performance of various forecasting models. The forecasts, generated in the initial project phase, are structured into matrices tailored for annual, quarterly, and monthly datasets. These matrices encompass predictions produced by diverse models, incorporating lagged values, predictor variables, and distinct selection methods. The Diebold-Mariano Test is utilized to compare all models, with the alternative set to "less" in R. For any two models, A and B, the hypotheses are formulated as follows:

Null Hypothesis (H_0): There is no difference in forecast accuracy between any two models.

$$H_0 : MFE_A = MFE_B$$

Alternative Hypothesis (H_1): The forecast accuracy of model A is better than the forecast accuracy of model B.

$$H_1 : MFE_A < MFE_B$$

The significance level is established at 0.05.

Figure 8 illustrates a heatmap of p-values for all pairs of models in the annual dataset. Blue boxes highlight p-values less than 0.05. In our test methodology, the horizontal axis represents models A, and the vertical axis represents models B. Notably, the linear model with lagged excess returns and lagged Consumption wealth income ratio outperform and demonstrate significant superiority over all other models. Following closely, the lagged Inflation emerges as the second-best performer, exhibiting noteworthy superiority over alternative models. Conversely, the linear model with all predictors displays the least forecasting accuracy, evident from p-values exceeding 0.05 in comparisons with other models.

Figure 9 presents a heatmap of p-values for all pairs of models in the quarterly dataset. Green boxes indicate p-values less than 0.05 and the horizontal axis represents models A, and the vertical axis represents models B. Once again, the lagged Consumption Wealth Income ratio excels and is significantly better than all other models. However, in this context, the excess returns lag 6 follows as the second-best performer, showcasing significant superiority over alternative models, while excess returns lag 8 secures the third-best position. In contrast, once again, the linear model with all predictors performs least effectively, evidenced by p-values greater than 0.05 in comparisons with other models.

Figure 10 illustrates a heatmap of p-values for all pairs of models in the monthly dataset, with pink boxes indicating p-values less than 0.05, while the horizontal axis represents models A, and the vertical axis represents models B. The lagged Cross Sectional beta Premium excels and is significantly better than all other models. Following closely, the excess returns lag 5 secures the second-best position, showcasing significant superiority over alternative models. the linear model with lagged excess return values and lagged Inflation model are at third place. In contrast, the regression tree with lagged excess return values performs least effectively, evidenced by p-values greater than 0.05 in comparisons with other models.

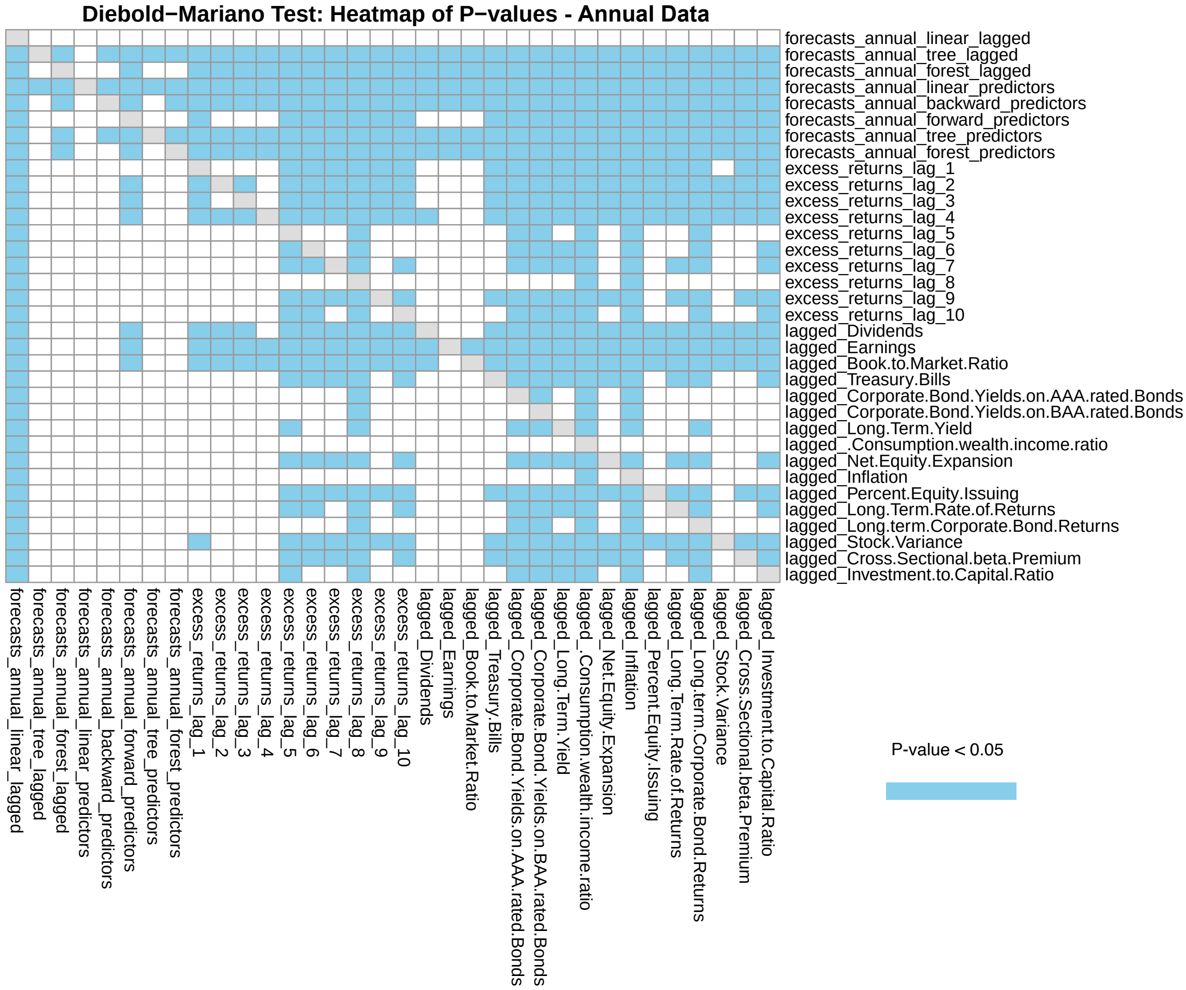

Figure 8: Diebold-Mariano Test Heatmap of P-values (Annual Data). The heatmap illustrates the pairwise statistical significance between different forecasting models based on the Diebold-Mariano test. Cells are color-coded to indicate significant (blue) and non-significant (white) differences in forecast performance. The significance threshold is set at 0.05. The horizontal axis represents models A, and the vertical axis represents models B.

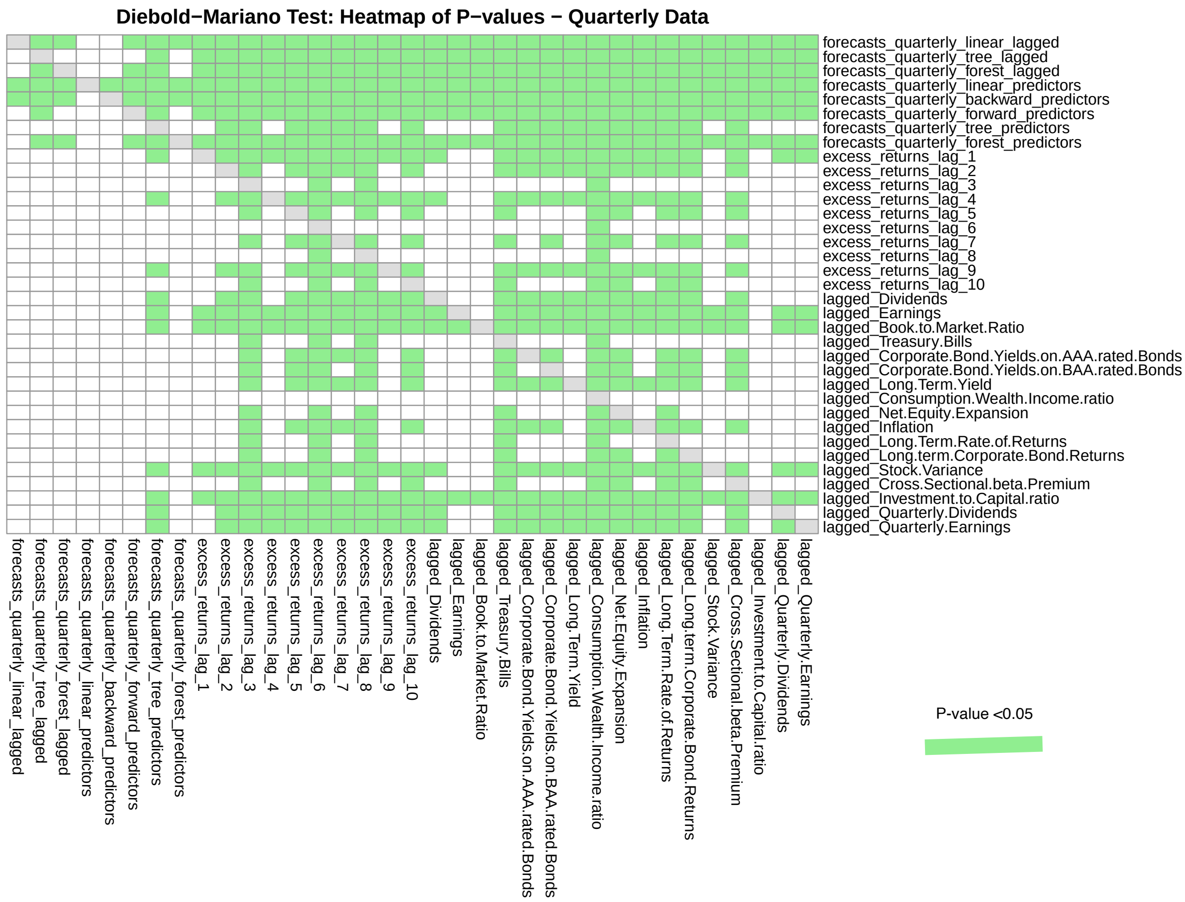

Figure 9: Diebold-Mariano Test Heatmap of P-values (Quarterly Data). The heatmap illustrates the pairwise statistical significance between different forecasting models based on the Diebold-Mariano test. Cells are color-coded to indicate significant (green) and non-significant (white) differences in forecast performance. The significance threshold is set at 0.05. The horizontal axis represents models A, and the vertical axis represents models B.

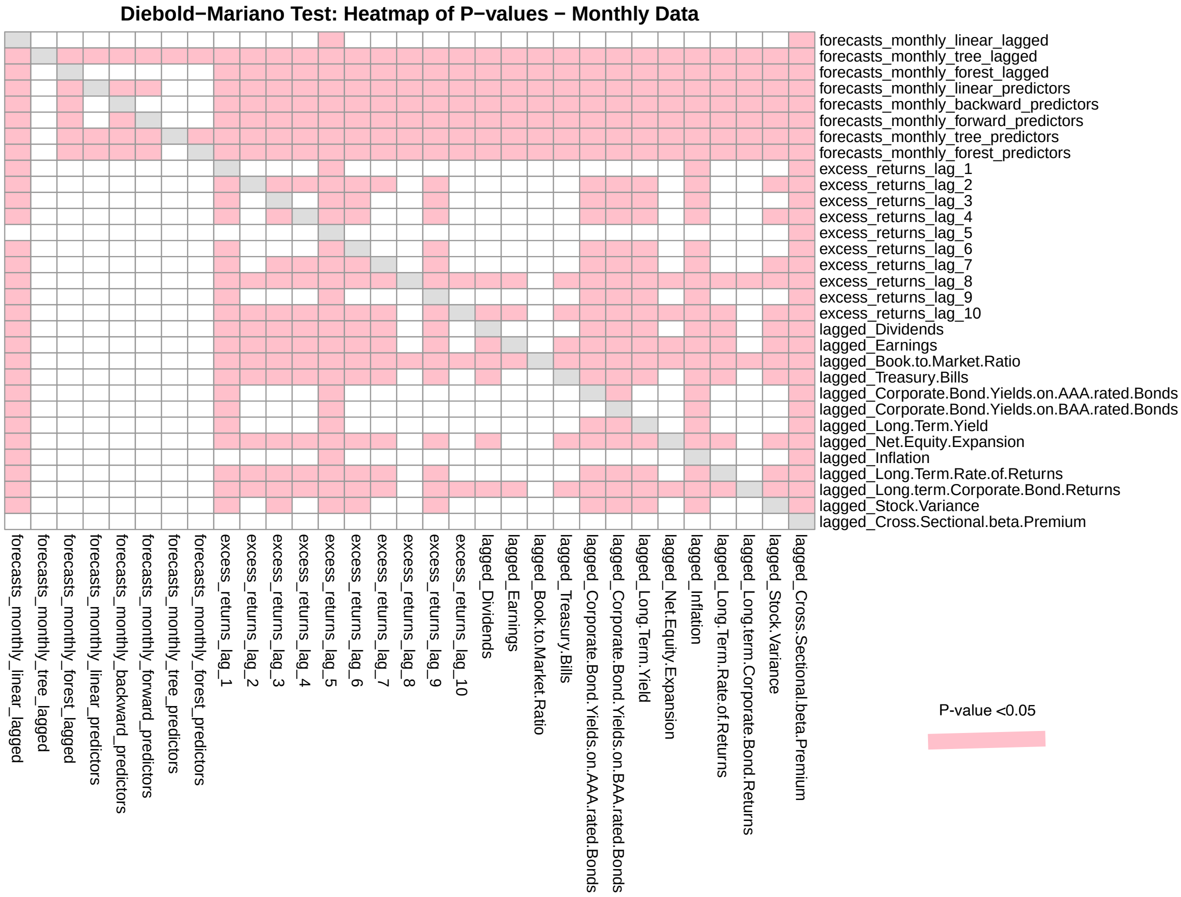

Figure 10: Diebold-Mariano Test Heatmap of P-values (Monthly Data). The heatmap illustrates the pairwise statistical significance between different forecasting models based on the Diebold-Mariano test. Cells are color-coded to indicate significant (pink) and non-significant (white) differences in forecast performance. The significance threshold is set at 0.05. The horizontal axis represents models A, and the vertical axis represents models B.

4.4 Evaluation and Reflection: Forecasting Models' Performance, Limitations, and Future Directions

The Diebold-Mariano Test results provide valuable insights into the forecasting ability of the models employed in our study. Notably, the linear model with lagged excess returns consistently outperformed other models across different data frequencies—annual, quarterly, and monthly. This suggests that the inclusion of set of predictors such as Consumption Wealth Income Ratio and Inflation contributes significantly to improved forecasting accuracy.

Conversely, the linear model with all predictors exhibits shortcomings, consistently displaying lower forecasting accuracy. This indicates limitations in capturing complex patterns within the data,

4.4.1 Shortcomings of the Models

Linear Model with Lagged Excess Returns: The linear model relying solely on lagged excess returns may be too simplistic, lacking the ability to capture intricate temporal dependencies in financial time series data.

Covariate Inclusion Models: Models incorporating covariates alongside lagged excess returns may face challenges in selecting and incorporating relevant variables. The performance may be contingent on the choice and relevance of covariates.

Regression Tree Model: Regression Trees are prone to overfitting, capturing noise in the training data and leading to poor generalization. Small changes in the data can result in significantly different tree structures, making the model less stable. Single trees might struggle to capture complex relationships present in the data compared to more sophisticated models.

Random Forest Model: Random Forests, especially with a large number of trees, can be computationally intensive during training. While Random Forests are robust against overfitting, there's still a possibility, particularly when the number of trees is too high. The ensemble nature of Random Forests may compromise the interpretability of the model compared to simpler models like linear regression.

4.4.2 Potential Improvements

Linear Model with Lagged Excess Returns: Experimenting with alternative lag structures or incorporating additional features may enhance the linear model's forecasting accuracy. This involves assessing different lag lengths and considering the inclusion of external macroeconomic indicators.

Covariate Inclusion Models: Implementing systematic approaches for covariate selection, such as feature engineering, dimensionality reduction, or model selection algorithms, can improve the reliability and interpretability of models incorporating covariates.

Regression Tree Model: Applying pruning techniques can help control the size of the tree and mitigate overfitting. Combining multiple Regression Trees into an ensemble, similar to Random Forest, can enhance predictive performance. Standardizing or normalizing input features can improve the stability and generalization of Regression Trees.

Random Forest Model: Careful tuning of hyperparameters, such as the number of trees and depth, can optimize Random Forest performance. Enhancing feature engineering or selecting more relevant features can improve Random Forest's ability to capture patterns. Leveraging parallel processing capabilities can alleviate computational intensity concerns, speeding up model training.

4.4.3 Further Outlook and Future Research

Advanced Machine Learning Techniques: Future research could explore advanced machine learning techniques, such as Long Short-Term Memory (LSTM) Networks, Gated Recurrent Units (GRU) or other deep learning to capture intricate patterns and nonlinear relationships within financial time series data. These techniques may offer improved forecasting accuracy in the presence of complex market dynamics.

Dynamic Model Specification: Models that adapt to changing market conditions through adaptive learning mechanisms, such as Kalman Filters and Adaptive Exponential Smoothing (AES), or through the incorporation of time-varying parameters, including Bayesian Structural Time Series (BSTS) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) for modeling time-varying volatility, have the potential to enhance forecasting accuracy. These approaches enable the models to capture evolving patterns in excess returns, making them more responsive and robust to dynamic market conditions.

Model Robustness Testing: Evaluating model robustness under various market scenarios, including periods of high volatility or economic downturns, is essential. Stress testing the forecasting models ensures their reliability across different market conditions.

5 Summary

In the ever-changing landscape of financial markets, accurately predicting stock returns proved to be a persistent challenge in our study. We delved into the equity premium, a crucial metric reflecting the anticipated excess return from stock market investments over the risk-free rate.

Expanding on prior projects, our analysis was enhanced by employing rolling-windows estimation and comparing forecasts from both linear and machine learning approaches. The goal was to gain insights into the predictability of excess returns and identify trends within financial data. Our methodology adopted a dual-dimensional approach: lagged-predictor-only and covariate-enriched forecasting. Real-time model updates and a maximal lag of 10 ensured adaptability to market dynamics.

In our extensive research on forecasting excess returns in financial markets, various models and evaluation metrics were utilized to assess their performance. The Root Mean Squared Forecast Error (RMSFE) played a pivotal role in measuring the accuracy of these models across different temporal scales. In one-step-ahead forecasting using lagged excess returns, the Linear Model emerged as the most accurate in the annual dataset, boasting the lowest RMSFE of 0.193. For the quarterly dataset, the Regression Tree model achieved the lowest RMSFE at 0.087, closely followed by Random Forest (0.088). In the monthly data, the Linear Model once again outperformed, presenting the lowest RMSFE at 0.044, compared to Random Forest (0.045) and Regression Tree (0.052) models.

The enhanced one-step-ahead forecasting with covariates refined the accuracy of the models. In annual data, variables such as the Consumption Wealth Income Ratio (0.1976) and lagged Inflation (0.1978) stood out for their relatively lower RMSFE values among individual lagged predictors. The Linear Predictors model recorded the highest RMSFE at 0.3711. In the quarterly dataset, lagged Consumption Wealth Income Ratio stood out with the lowest RMSFE of 0.0812, while the Regression Tree Predictors model showcased a better RMSFE at 0.0853. Linear Predictors, with an RMSFE of 0.3214, demonstrated a higher error. For monthly data, all predictor categories consistently exhibited low RMSFE values, with Linear Predictors at 0.0464, Backward Predictors at 0.0457, and Forward Predictors at 0.0457, while the lowest RMSFE was observed for lagged Inflation (0.0444).

The Diebold-Mariano Test, a crucial tool for model comparison, identified the linear model with lagged excess returns and lagged Consumption Wealth Income Ratio as significantly outperforming other models in annual data. In the quarterly dataset, the lagged Consumption Wealth Income Ratio excelled, displaying significant superiority. Excess returns lag 6 and lag 8 also performed exceptionally well. In monthly data, the lagged Cross Sectional beta Premium demonstrated remarkable performance, securing the top position. Excess returns lag 5 also exhibited significant superiority.

In summary, the Linear Model with lagged excess returns consistently stood out across different temporal granularities, showcasing the lowest RMSFE values in both one-step-ahead forecasting and the Diebold-Mariano Test. Its adeptness in incorporating lagged information and key covariates, particularly the Consumption Wealth Income Ratio, positioned it as a robust choice for forecasting excess returns in financial markets.

Additionally, deeper understanding and continuous refinement of forecasting models for excess returns in financial markets, future research avenues could explore the synergies of ensemble methods, unravel the intricacies of feature importance, and delve into the dynamic nature of time-varying predictors. Further optimization through hyperparameter tuning and robustness testing against varied datasets and market conditions can enhance the reliability of predictive models. The inclusion of volatility models, examination of the impact of economic events, and extension of forecasting horizons beyond one step ahead are vital directions for a more comprehensive analysis.

Bibliography

Francis X. Diebold and Roberto S. Mariano. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13:253–263, 1995.

Amit Goyal. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, 2022. URL `https://sites.google.com/view/agoyal145`.

Rob Hyndman, George Athanasopoulos, Christoph Bergmeir, Gabriel Caceres, Leanne Chhay, Mitchell O'Hara-Wild, Fotios Petropoulos, Slava Razbash, Earo Wang, and Farah Yasmeen. *forecast: Forecasting functions for time series and linear models*, 2023. URL `https://pkg.robjhyndman.com/forecast/`. R package version 8.21.1.

Atsushi Inoue, Lu Jin, and Barbara Rossi. Rolling window selection for out-of-sample forecasting with time-varying parameters. 2016. URL `https://ncbi.nlm.nih.gov/pmc/articles/PMC2770192/`.

Raky Julio, Andres Monzon, and Yusak O. Susilo. Identifying key elements for user satisfaction of bike-sharing systems: a combination of direct and indirect evaluations. *Journal Name*, 2022. doi: 10.1007/s11116-022-10335-3. URL `https://link.springer.com/article/10.1007/s11116-022-10335-3`.

Alboukadel Kassambara. *ggpubr: ggplot2' Based Publication Ready Plots*, 2022. URL `https://rpkgs.datanovia.com/ggpubr/`. R package version 3.4.

Raivo Kolde. *pheatmap: Pretty Heatmaps*, 2019. R package version 1.0.12.

Max Kuhn and Others. *caret: Classification and Regression Training*, 2022. URL `https://CRAN.R-project.org/package=caret`. R package version 6.0-91.

Andy Liaw and Matthew Wiener. *randomForest: Breiman and Cutler's Random Forests for Classification and Regression*, 2022. URL `https://CRAN.R-project.org/package=randomForest`. R package version 4.6-16.

Stephen Milborrow. *rpart.plot: Plot 'rpart' Models: An Enhanced Version of 'plot.rpart'*, 2022. URL `https://CRAN.R-project.org/package=rpart.plot`. R package version 3.1.5.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL `https://www.R-project.org/`.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023. URL https://www.R-project.org/.

Terry Therneau and Beth Atkinson. *rpart: Recursive Partitioning and Regression Trees*, 2022. URL https://CRAN.R-project.org/package=rpart. R package version 4.1-16.

Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*, 2016. URL https://ggplot2.tidyverse.org.

Hadley Wickham, Romain François, Lionel Henry, and Kirill Müller. *dplyr: A Grammar of Data Manipulation*, 2022. https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr.

Hadley Wickham, Jim Hester, and Jennifer Bryan. *readr: Read Rectangular Text Data*, 2023a. URL https://readr.tidyverse.org. R package version 2.1.4, https://github.com/tidyverse/readr.

Hadley Wickham, Davis Vaughan, and Maximilian Girlich. *tidyr: Tidy Messy Data*, 2023b. URL https://tidyr.tidyverse.org. R package version 1.3.0, https://github.com/tidyverse/tidyr.