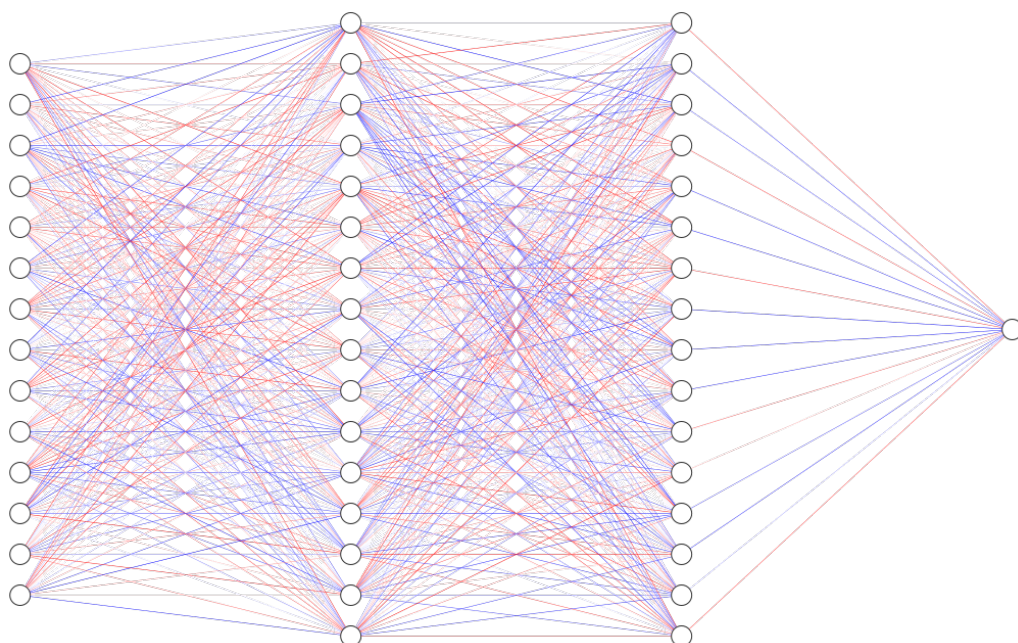


**ERASMUS UNIVERSITY ROTTERDAM**  
**ERASMUS SCHOOL OF ECONOMICS**  
**MSc Economics & Business**  
**Specialization Financial Economics**

# **Hybrid Portfolio Strategy: PCA-Fundamental Sorting and PCA-LSTM Neural Network for Technical Indicators**

**Evidence from S&P Europe 350 Index Constituents**



**Author:** Kerem Koruk  
**Student number:** 578245  
**Thesis supervisor:** Dr. Jan Lemmen  
**Second reader:** Dr. Ruben de Blik  
**Finish date:** 30 July 2024

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

## **ABSTRACT**

In this thesis, I examined the performance of a hybrid portfolio strategy integrating principal component analysis (PCA) and Long Short-Term Memory (LSTM) networks to achieve consistent risk-adjusted abnormal return using the investible universe of the S&P Europe 350 Index constituents. By developing a monthly balance sheet score index and daily price index from security-level fundamental and technical data, this study constructs a monthly rebalanced equal-weighted long-short portfolio free of survivorship and look-ahead biases. This hybrid strategy significantly outperforms the index, momentum factor, and 5-factor Fama-French model with an average annualized Sharpe ratio of 1.45, Information ratio of 1.26, and (compounded) average alpha of 0.38, while failing to offset the positive systematic risk with an average annualized Treynor ratio of -0.57. Notwithstanding, the hybrid strategy experienced a maximum drawdown of -16.47% from 2019 to 2024, implying the strategy's stability against market volatility.

**Keywords:** Portfolio Management, Long Short-Term Memory, Quantitative Investing, Principal Component Analysis, European Equity Market

**JEL Classification:** G11, G12

# TABLE OF CONTENTS

|   |     |
|---|-----|
| ABSTRACT .....  | ii  |
| TABLE OF CONTENTS .....                                       | iii |
| CHAPTER 1 Introduction .....                                  | 1   |
| CHAPTER 2 Theoretical Framework .....                         | 4   |
| 2.1 Quantitative Investment Methods .....                     | 4   |
| 2.1.1. Efficiency of Capital Markets .....                    | 5   |
| 2.1.2. Portfolio Management.....                              | 6   |
| 2.1.3. Empirical Asset Pricing Models.....                    | 7   |
| 2.1.4. Modern Quantitative Investment Methods.....            | 9   |
| 2.2. Deep Learning.....                                       | 10  |
| 2.2.1. Long Short-Term Memory Neural Networks .....           | 11  |
| 2.3. Hypotheses .....   | 12  |
| CHAPTER 3 Data .....  | 14  |
| 3.1. Historical Constituents of the S&P Europe 350 Index..... | 14  |
| 3.2. Fundamental Data .....                                   | 16  |
| 3.3. Technical Data .....                                     | 17  |
| 3.4. External Data.....                                       | 18  |
| 3.5. Data Cleaning Process.....                               | 19  |
| 3.6. Summary Statistics.....                                  | 20  |
| CHAPTER 4 Method .....  | 25  |
| 4.1. Principal Component Analysis (PCA) .....                 | 29  |
| 4.2. LSTM Neural Network .....                                | 30  |
| 4.3. Portfolio Performance Metrics.....                       | 33  |
| CHAPTER 5 Results & Discussion .....                          | 35  |
| 5.1. Fundamental Data .....                                   | 35  |
| 5.1.1 Validity Test.....                                      | 35  |
| 5.1.2 Stability Test .....                                    | 36  |
| 5.1.3 PCA Analysis .....                                      | 37  |
| 5.2. Technical Data .....                                     | 39  |
| 5.2.1 PCA Analysis .....                                      | 39  |
| 5.2.2 LSTM Neural Network .....                               | 41  |
| 5.3. Portfolio Performance .....                              | 43  |
| CHAPTER 6 Conclusion.....                                     | 49  |
| REFERENCES .....  | 51  |
| APPENDIX A Data.....  | 56  |
| APPENDIX B PCA Analysis.....                                  | 58  |
| APPENDIX C Alpha Regressions.....                             | 62  |



## CHAPTER 1 Introduction

Since the existence of equity markets, practitioners and extensive academic literature have strived to predict future returns using historical prices as a predictor, identifying factors like price momentum and reversal as reliable predictors. The recent improvement in the understanding of how human behavior and psychology dictate price dynamics (Barberis, 2018) implies that prices might have complicated patterns that are difficult to unravel using traditional tools for empirical finance. In this study, I re-examine price movements from a deep learning standpoint, considering the never-ending curiosity the market has for an explanation for the behavior of financial assets or financial purposes. More specifically, Long Short-Term Memory (LSTM), a type of recurrent neural network (RNN), is among the best-suited deep learning techniques due to its capacity to efficiently deal with sequential data while predicting market movements (Sezer et al., 2020). However, transitioning to a deep learning framework for stock price-based return prediction poses two significant challenges: establishing a methodology that offers a balance between the ability to detect intricate and nonlinear patterns and consistency with established asset pricing models to have an economic rationale for the financial forecasting of deep learning techniques.

Recently, with the growing insights from deep learning methods, researchers, in particular in the fields of Quantitative Finance and Computer Science, have been focusing on whether and how deep learning techniques can have significant predictive power and generate abnormal returns in the financial market despite the lack of search for the economic rationale behind their findings. The importance of their studies originates from the contradictory relationship of promising financial time-series forecasting of deep learning techniques with any form of Efficient Market Hypothesis. For instance, Jiang et al. (2023) study the predictive power of image-based convolutional neural networks (CNN), a different type of RNN, on stock price trend movements based on the U.S. stock returns with data spanning from 1993 to 2019. They find statistically significant returns and strong Sharpe ratios (even when benchmarked against other common price trend strategies) from High-Low portfolios (10-1 deciles) for both short- and long-term horizons, where short-term horizon trading yields greater returns. The paper also refers to the LSTM model with its possibility of outperforming CNN in terms of a more accurate prediction model. The paper of Yan and Ouyang (2018) uses the LSTM technique to predict the following day's closing price of the Shanghai Stock Exchange 180 (SSE 180) from 2012 to 2017, where they find statistically significant predictive power of the LSTM model for the SSE 180 index over other machine learning models such as support vector machine and multilayer perceptron. Likewise, the papers of Yang (2021) and Liu et al. (2017) use LSTM models to predict financial indicators in the U.S. market and stock prices of the Chinese stock market, respectively. These papers also conclude the LSTM models achieve significant accuracy in time-series prediction and outperformance over the benchmark indexes.

Most studies that focus on deep-learning methods for financial time-series forecasting use different investment frequencies, from weekly to quarterly, to investigate whether the window size for the investment strategy significantly affects the performance of a portfolio. However, the academic literature lacks adopting a rolling window during portfolio rebalancing despite its further improvement in the accuracy of predictions as recursively retraining the models (while rebalancing the portfolio) is primarily subjected to the capacity limitation of computational resources (Jiang et al., 2023). Furthermore, it is important to note that the deep learning models in the literature typically use either technical analysis, which focuses on historical price and volume data, or fundamental analysis, which relies on financial statements and economic indicators, to inform their predictions and strategies (Heaton et al., 2016; Sezer et. al, 2020). However, Bettman et al. (2009) show that technical and fundamental analyses are implied not to substitute but to complement each other to improve investment decision-making. Moreover, most studies examining the financial time-series forecasting power of deep learning models mainly focus on the U.S. and Asia Pacific markets, with limited attention given to the European market as a whole. Therefore, in this thesis, I aim to answer the following research question: *How does the hybrid portfolio strategy, using fundamental sorting and LSTM Networks for technical indicators, perform relative to the S&P Europe 350 Index, momentum factor, and 5-factor Fama-French model?*

In this study, I developed a monthly rebalanced hybrid equal-weighted (EW) long-short hybrid portfolio strategy that integrates principal component analysis (PCA) both for fundamental factor sorting (optionally) and for Long Short-Term Memory (LSTM) neural networks, which derives trading signals from technical indicators, using the constituents of the S&P Europe 350 Index. Thus, I first screened the initially collected 26 fundamental data variables based on their validity and stability through for which information coefficient (IC) and information ratio (IR) are used, respectively, during the in-sample period. The fundamental data variables that passed both the validity and stability tests are chosen as core fundamentals. After the correlation matrix check of core fundamentals, I determined to employ principal component analysis (PCA) to prevent multicollinearity and reduce dimensions to create a balance sheet score index on a company level on a monthly basis for the out-of-sample period. Subsequently, I create monthly equal-weighted quintiles sorted on these balance sheet score indexes, in which the first quintile Q1 (fifth quintile Q5) represents the worst-performing (best-forming) securities, for the out-of-sample period. As for the technical data section, the initial 6 daily historical trading volume and price data variables were processed through principal component analysis (PCA) to reduce dimensions in order to create a daily price index on a company level for the in-sample period. After creating such a daily price index, the LSTM neural network was employed to train and validate the daily price indexes for the in-sample period to provide out-of-sample daily price index forecasts. Using the LSTM networks, I aimed at identifying stocks with the strongest increasing and decreasing signals each month within the highest and lowest fundamental quintiles, respectively. Given the monthly equal-weighted quintiles sorted on balance sheet score indexes and incorporating trading signals of the LSTM network based on the daily price index, I constructed the final monthly rebalanced EW long-short hybrid portfolio in which I took

a long position in the top 10% of Q5 (best-performing quintile) securities, while I take a short position in the bottom 10% of Q1 (worst-performing quintile) securities both based on the forecasted monthly return of daily price index in the out-of-sample period. The performance of this hybrid portfolio strategy will be benchmarked against a traditional momentum factor, factors of the 5-factor Fama-French model, and the S&P Europe 350 index. The sample size will vary from 350 to 450 companies over the study period. Data sources will include the Bloomberg Terminal for historical constituents of the S&P Europe 350 Index, the Thomson/Refinitiv database for monthly historical fundamental values, and Yahoo Finance for daily historical technical indicators. This study aims to determine whether the hybrid strategy can generate consistent abnormal returns compared to market benchmarks.

Prior to conducting these analyses, I hypothesize that the monthly rebalanced EW long-short portfolios generated by the hybrid strategy of fundamental factor sorting and LSTM network will significantly outperform both the S&P Europe 350 Index and the momentum factor, consistent with the findings of Jiang et al. (2023). Additionally, I expect that this hybrid strategy will outperform each factor of the 5-factor Fama-French model. The underlying reason of this hypothesis is similar to the hypothesis of outperforming momentum factor. I expect that by incorporating valid and stable balance sheet materials and technical indicators into the hybrid portfolio strategy can potentially allow outperforming relatively more restricted individual factors of the 5-factor model. Consequently, as this paper incorporates broader fundamental data pool relative to the 5-factor model along with 6 daily historical trading volume and price data variables, I hypothesize that the hybrid portfolio strategy through its holistic approach can outperform the 5-factor Fama-French model as a whole. Furthermore, I predict that constructing portfolios by selecting stocks from various European stock markets, without restricting the relatively narrow investable universe to S&P 350 Europe Index constituents, will achieve higher diversification and ultimately improve the performance of LSTM-based portfolios.

The rest of this thesis is structured as follows. Chapter 2 follows next and provides the theoretical framework, beginning with a thorough review of quantitative investment methods. This includes discussions on the efficiency of capital markets, portfolio management, empirical asset pricing models, and modern quantitative investment techniques. Followingly, the chapter delves into deep learning methods, with a particular focus on Long Short-Term Memory (LSTM) neural networks. Chapter 3 describes the sample dataset, including the data collection procedure and key summary statistics. Chapter 4 outlines the statistical methodologies used in this study. Chapters 5 and 6 are dedicated to presenting the results and concluding remarks, respectively.

## **CHAPTER 2 Theoretical Framework**

The rapid expansion of financial markets has significantly broadened the scope of investment opportunities available to investors, leading to diverse investment strategies, each with its unique approach. These strategies mainly include, but are not limited to, trend investment, relying on identifying patterns in historical market data; value investment, involving analyzing fundamental factors such as company earnings and assets; and quantitative investment, utilizing advanced computer technology and mathematical models.

Of these approaches, quantitative investment has grown in favor over the last decade, largely due to its rigorous methodology, enabling investors to make informed decisions based on a thorough analysis of data. Furthermore, renowned for their agility in adjusting relatively faster to market shifts, quantitative investment strategies allow investors to capitalize on opportunities as they arise, making quantitative investment increasingly attractive for efficient active portfolio management.

Alongside the ongoing advancements in computer technology, deep learning has emerged as a potent instrument for processing complex, nonlinear time series data and improving the accuracy of predicting stock prices. By integrating quantitative investment strategies, both with sophisticated stock price prediction models and value investment strategies, there is a potential to achieve consistent outperformance over the market, which directly challenges the efficient market hypothesis (EMH) of Fama (1970). In this research study, I will aim to investigate three primary dimensions: quantitative investment strategies, stock price prediction, and deep learning techniques focusing on long short-term memory (LSTM) neural network.

### ***2.1 Quantitative Investment Methods***

Quantitative investment methods include using precise mathematical, statistical, and economic theory, integrated with advanced information technology, to examine the financial markets and to yield consistent abnormal returns over the market through constructing quantitative investment-based portfolios (Chandrashekar, 2015; Ma, 2020). Having a rigorous historical academic background and experiencing fluctuations in popularity among investors, these methods are ‘time-tested approaches’ based on robust economic principles, extensive academic research, and practical analysis (Chandrashekar, 2015; Fabozzi, 2010). Their ability to allow quick portfolio trading and active portfolio management and to identify and optimize multiple investment prospects increase the unbiased, efficiency, and transparency in investor decision-making (Litterman, 2003). The rapid advancements in big data and computing power technologies since 2010 have been transforming the financial sector (American Economic Association, 2020), and quantitative investment methods have begun replacing traditional investment methods, particularly in hedged funds (Farboodi and Veldkamp, 2020).

The theoretical basis of the quantitative investment methods I investigate in this paper stems from the EMH and modern portfolio theory (MPT) and primarily aims to scrutinize and challenge these

findings. Extensive academic research has been, for a long time, examining the concept of informational efficiency across capital markets and MPT and investigating whether active portfolio management could consistently yield abnormal returns (Jacobsen, 1988). A portfolio achieving such performance is considered to outperform the market, thereby creating extra value for investors. However, the empirical evidence in the literature still indicates no consensus in this regard, partly due to changing economic and market conditions (Tıtan, 2015; Naseer et al., 2015).

### **2.1.1. Efficiency of Capital Markets**

Fama (1965) introduces the first concept of efficiency in which the author asserts that investors, regardless of their respective skill or knowledge, cannot achieve abnormal profits over the market performance unless they invest in riskier assets Samuelson (2016). Followingly, Fama (1970) elaborates on the definition of the hypothesis by identifying three types of efficiency: weak, semi-strong, and strong. He defines an efficient market in which a large number of homo economicus compete with each other to predict price changes of securities where all significant information regarding securities is readily accessible to all market participants. The weak form suggests that the historical data is already factored in security prices, making past prices and volumes irrelevant for predicting the future value of securities. Building on this, the semi-strong form specifies that all publicly accessible information of companies, such as financial statements and news, are fully priced in security prices, additionally rendering the fundamental analysis ineffective in predicting the future value of securities. Lastly, the strong form states that all public and private information of companies is completely reflected in security prices, indicating even insider information doesn't provide investors an edge in achieving abnormal returns.

The EMH brought forth prominent theories including the fair game, sub-martingale, and random walk models. The first theory implies the demand-supply equilibrium for the security price is determined by the expected return, which is shaped by securities' associated risk in the same manner as in Sharpe (1964) and Lintner's (1965) capital asset pricing model (CAPM) where investors have returns that are consistent with the risk- associated when all information available is factored in securities' prices. To further the fair game model, the sub-martingale theory indicates that future prices are expected to exceed current prices; hence, trading strategies based on historical data cannot outperform buy-and-hold strategies on a risk-adjusted basis (Fama, 1970). Likewise, the random walk hypothesis, as defined by Osborne (1959) and Alexander (1961), implies securities' price fluctuations are independent and identically distributed, rendering historical trends ineffective for forecasting securities' price movements.

Malkiel (2003), who is also a proponent of the random walk hypothesis, further augments the EMH by asserting that inexperienced or uninformed investors with diverse portfolios can achieve profits comparable to the ones achieved by financial experts. Furthermore, despite acknowledging the

persistence of mispricings in the market, Malkiel (2003) addresses that these mispricings often fail to offset transaction costs for investors to generate positive abnormal returns.

In this paper, I strive to challenge the weak and semi-strong forms of EMH by primarily using historical data (including technical indicators) and fundamental factors of companies to construct portfolios with the aim of generating consistently significantly positive abnormal returns over market performance. Despite this, it is worth mentioning that investment strategies that claim to consistently beat the market are viewed - and should be - with suspicion due to the aforementioned assumptions of EMH (Malkiel, 2021).

### **2.1.2. Portfolio Management**

Given the theories of the EMH, particularly the fair game model, optimizing asset allocation is one of the primary aspects of making investment decisions amidst uncertainty and risk. As the backbone of modern quantitative investing, Roy (1952) and Markowitz (1952) introduce the mathematical derivation of diversification in which Markowitz (1952) pioneered the development of the modern portfolio theory (MPT) that highlights diversification in a mean-variance framework for investment decision-making (Sharpe, 1992). Based on the MPT, diversification formalizes the wisdom of not concentrating all your assets in one place. Recognizing the unpredictable swings in securities' prices and the risk of large losses due to volatility, Markowitz (1952) proposes using standard deviation to assess investment uncertainty, making portfolio optimization mathematically feasible. Given the findings of the MPT, Markowitz (1952) demonstrates how to construct efficient portfolios that lie on the efficient frontier, which, in quantitative investing, represent the set of portfolios that offer the highest expected return for a given level of risk, and vice versa. For instance, with predictions of stock returns, their standard deviations, and their correlation, investors can construct efficient portfolios and reach the efficient frontier of the mean-variance framework.

In the field of portfolio management, understanding the expected performance and value of assets is crucial. Analysts usually look for assets to add to their long positions, partly due to short-selling and transaction cost-related restrictions, but some (Jacobs et al., 1999; Bender et al., 2010; Blackrock, 2024) also suggest using long-short portfolios to potentially boost returns. By incorporating a short-selling position into the portfolio strategy, investors increase their flexibility and ability to profit from both rising and falling stock prices (Jacobs et al, 1993). Furthermore, Jacobs et al. (1999) highlight that the real benefits of employing a short strategy are best observed when the portfolio is viewed and organized as a cohesive whole incorporating both long and short positions. A short strategy is more than just merging two distinct portfolios; it represents a unified approach where the selection of long and short positions is interconnected. Moreover, Blackrock (2024) underlines that long-short portfolios can help reduce the impact of market risks and lower overall portfolio volatility. These approaches are less dependent on market uptrends, providing opportunities for profits regardless of price movements and leading to the possibility of having a more diversified portfolio that is less correlated to equity markets.

Similarly, Jacobs and Levy (1996) emphasize that unlike long-only strategies, which are limited by benchmark weights, long-short portfolios offer the adaptability to overweight or underweight securities. Based on the 3rd myth of Jacobs and Levy (1996), long-short strategies are not bound by the weights of the underlying index, and any sensitivity to the index is mitigated by balancing long and short positions.

In contrast to the favorable findings associated with the long-short strategy, Michaud (1993) cautions that its improved expected return is accompanied by heightened portfolio risk, potentially conflicting with the objectives of long-term investors given the transaction costs associated with the strategy. Aligned with the efficient frontier, investors should be aware of the potential of the increased level of risk with the long-short strategy to outperform the long-only strategy. Furthermore, Huij et al. (2014) indicate that despite the theoretical superiority of the long-short strategy, the actual risk-adjusted returns and real-world practices show a preference, for sticking to a long-only strategy.

### ***2.1.3. Empirical Asset Pricing Models***

Accurate pricing of each asset is crucial when it comes to portfolio creation and management as it serves as the foundation for achieving optimal asset allocation. Therefore, employing a suitable asset pricing model is of great importance in ensuring accurate results. As aforementioned in Markowitz's (1952) MPT, investors see a risky investment as attractive only if it promises a higher rate of return. In simpler terms, when compared to a safe investment that gives a return equal to the time value of money, a risky investment needs to offer investors an additional reward to make up for potential losses. MPT's introduction of the concept of risk, specifically regarding the standard deviation of returns, played a crucial role in the development of quantitative investing during the 1960s as it led to the creation of an influential economic model of the Capital Asset Pricing Model (CAPM) that established a strong connection between the risk of a security and its expected return — introduced by Sharpe (1964) and Lintner (1965), the CAPM functions based on the assumption that all securities are influenced by changes in the overall market. The beta of a security represents the portion of its risk that is attributed to market exposure. As per the CAPM, only this non-diversifiable risk component should be compensated by the market. Any other price fluctuations in security, known as diversifiable risk, become insignificant when an investor holds a diversified portfolio and therefore should not be rewarded by the market. Essentially, the CAPM states that the expected returns of individual securities, adjusted for the risk-free rate, should demonstrate a linear correlation with their betas. The line, illustrating this linear relationship between the expected returns and beta coefficients, is called the Security Market Line (SML), which acts as a benchmark for evaluating asset valuation. Similarly, the alpha coefficient, introduced by Jensen (1968), is a risk-adjusted performance measure that compares actual portfolio returns with those predicted by CAPM. In the CAPM formula, alpha is incorporated as the intercept, indicating whether an investment outperforms or underperforms relative to CAPM's prediction.

Over the past five decades, both scholars and practitioners have strived to demonstrate the inefficiencies present in financial markets. The underlying objective driving this study is

straightforward: if markets exhibit inefficiencies and anomalies, identifying and capitalizing on these anomalies could potentially enable one to outperform the market, at least in principle. The research of Fama and Macbeth (1973) on the empirical testing of the CAPM indicated the existence of factors, other than the market risk, that consistently affect the securities' expected returns. After the finding of Fama and Macbeth (1973), many research studies (e.g., Haugen and Heins, 1975; Basu, 1977; Stattman, 1980; Banz, 1981; Rosenberg et al., 1985) have provided compelling evidence that the differences in expected returns among securities are not significantly related to beta, but instead show a strong correlation with various other financial characteristics. The majority of the findings in the literature suggest that while there is only a limited variation in expected returns based on beta, there is a significant variation across these alternative financial characteristics-based dimensions. These studies have analyzed the volatility in expected returns from multiple perspectives: beta and another financial characteristic(s). Haugen and Heins (1975) examine whether firm-level characteristics, such as price-to-earnings ratios, dividend yields, and leverage, explain security returns better than the CAPM beta. Similarly, Basu (1977) presents that value stocks with low price-to-earnings ratios yield, on average, higher risk-adjusted returns compared to growth stocks. Banz (1981) found the existence of the size anomaly, where firms with small market capitalization yield, on average, higher risk-adjusted returns compared to those with large market capitalization. Moreover, Stattman (1980) and Rosenberg et al. (1985) propose the value anomaly, in which stocks with high book-to-price ratios yield, on average, higher risk-adjusted returns than those with low book-to-price ratios. Aside from the firm-level financial characteristics-based anomalies, Jegadeesh and Titman (1993) and Assness et al. (1997) suggest a trend-related anomaly, called momentum, where “winner” stocks with higher past returns yield higher risk-adjusted returns than “loser” stocks with lower past returns. These studies and many more in the literature hint at the compelling proof of market inefficiency.

These findings of anomalies in the market have subsequently led to the construction of numerous asset pricing models in the literature. Fama and French (1992) improved the explanatory power of the CAPM by incorporating size and value factors into the model, constructing the renowned three-factor model. However, despite these improvements, Fama and French (1996) self-criticized their three-factor model as lacking in explaining the variation in returns, particularly among momentum portfolios. This was a particularly important finding as the persistence of the momentum effect has been observed across different time frames and countries Assness et al. (1997). In this regard, Carhart (1997) constructs a four-factor model, building on the three-factor model, by incorporating the momentum factor and found the momentum factor capturing significant variation in equity mutual fund returns. Furthermore, Fama and French (2015) introduce a five-factor model by incorporating investment and profitability factors into the three-factor model. In their model, the investment factor is the average return difference between the conservative firms with low investment levels and aggressive firms with high investment levels, whereas the profitability factor suggests the average return difference between robust profitable firms and weak profitable firms.

To describe the rapidly increasing amount of empirical factors in the literature, Cochrane (2011) coined the term “factor zoo”, for which Campbell et al. (2016) showed over 300 factors proposed in the literature. Despite the availability of a vast number of factors, Campbell et al. (2016) and Mclean and Pontiff (2016) underline the challenges that are observed in many factors in the factor zoo such as multiple testing, overfitting, and lack of out-of-sample performance and also highlight the possibility of these factors not persisting over time and disappearing completely. The literature widely acknowledges that the market is not perfect and there are various factors that can account for the returns of securities. However, differentiating between valid and invalid factors presents a more intricate obstacle (Hou et al., 2020).

#### **2.1.4. Modern Quantitative Investment Methods**

Quantitative investing has significantly grown since the Ph.D thesis of Bachelier (1964) on the fluctuation of security prices; academicians and practitioners have ever since tried to pursue diligent research-based methods to chase significant alpha and unravel systematic and idiosyncratic risks associated with their methods. Since the 2007-2008 global financial crisis, there has been growing institutional investor interest in quantitative methods (Farboodi and Veldkamp, 2020) due to their unbiased and systematic methodologies and time-efficient ability to deliver consistent risk-adjusted returns. Despite falling short of performance expectations during the global financial crisis, quantitative methods rebounded quickly and partly exceeded traditional investment methods based on fundamental analysis in terms of assets under management (Farboodi and Veldkamp, 2020).

As discussed previously, traditional investment approaches have included both the alpha and beta strategies. Stock-picking and timing strategies characterize alpha strategies since they seek excess returns. Active selection of particular securities based on analysis of fundamentals or technical data is aimed at discovering undervalued or overvalued assets that are expected to outperform the market. Market timing, another dimension of alpha strategies, involves strategic decisions regarding buy or sell positions based on forecasts of market movements to capitalize on short-term fluctuations. Besides, value investing, which is based on the theory of hedging, tries to separate alpha returns from beta returns by using derivative instruments to hedge against systematic risk and produce additional returns. In value investing strategy, the attempt is to pick out those undervalued stocks with solid fundamentals that usually have low price-to-earnings multiples and hold them until their market price accurately reflects their intrinsic worth. On the other hand, the beta strategy aims at building a portfolio that tends to precisely imitate the market by either reproducing the market portfolio or replicating its price movement. This is to capture the market returns passively. Typically, the beta strategies are in the form of index funds or ETFs that follow a specific market index.

At present, alpha strategies are highly preferred as a quantitative investment approach. These strategies can be broadly classified as, including but not limited to, factor investing (Hanauer, 2020), multi-factor investing (Clarke et al., 2016), alternative risk premia (Botosan and Plumlee, 2005; Ilmanen

et al, 2021), investor sentiment-based (Brown and Cliff, 2005), and quantitative systematic (Yin et al, 2021) investing methods. Among these, the quantitative systematic investing methods through incorporating multi-factor models are particularly notable for their comprehensive abilities to optimize complicated goals and restrictions by quickly capturing and adjusting changes in the market sentiment.

Particularly with the recent development in advanced information technologies such as big data processing technologies (Farboodi and Veldkamp, 2020) and explainable artificial intelligence (XAI) (Zhang et al., 2020), quantitative methods have begun adopting machine learning algorithms (Gu et al, 2018). These algorithms provide a variety of methodologies, such as support vector models, tree-based methods, deep learning models, and hybrid models (de Prado, 2020). The most essential feature of these algorithm models is to detect the anomalies or mispricings that have a strong association with stock returns based on past data to generate abnormal returns.

## **2.2. Deep Learning**

The finance sector has long been concerned with accurately forecasting financial time series (FTS) data. The literature on FTS identifies that machine learning (ML) techniques tend to perform better in this respect than traditional methods of time series forecasting, such as generalized autoregressive conditional heteroskedasticity (GARCH), autoregressive integrated moving average (ARIMA), and vector autoregression (VAR) (Bagnara, 2024). The conventional models are relatively more restricted due to their reliance on linear assumptions to predict FTS data. But in reality, stock prices are primarily nonlinear and depend on a couple of factors. Also, as the period of data collection increases, the noise level in data increases, hence making it more challenging for an accurate forecast of data for the traditional forecasting of time series methods. Due to the ability of machine learning models to handle the non-linearity and noise in the FTS data, automated trading systems are increasingly being used, particularly in algorithmic trading, while the search for abnormal returns continues to encourage developing better machine learning models.

Among ML techniques, deep learning (DL) techniques, a subset of artificial neural networks (ANN), have been highly favored by academics and practitioners in recent years due to their significant outperformance over traditional ML techniques in FTS forecasting in the literature (Li and Ma, 2010; Dixon et. al, 2017). A significant strength of the characteristic of multiple processing layers in DL models for data modeling is a high-level abstraction from raw financial time-series data (Sezer et al., 2020), and the primary strength of DL models lies in the capacity to automatically extract latent features of raw financial time-series data with a general-purpose learning algorithm (Sezer et al., 2020; Jiang et al., 2023). The following list of DL models that can be put to use in the literature focusing on forecasting financial time series (FTS) are as follows: Deep Multilayer Perceptron (DMLP), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Restricted Boltzmann Machines (RBMs), Deep Belief Networks (DBN), Autoencoder (AE), and Deep Reinforcement Learning (DRL).

While DL methods have shown promise, early attempts such as the back-propagation (BP) neural network model introduced by White (1988) to predict daily stock returns encountered significant limitations such as the model not explaining the whole data sequence period, difficulties in attaining the optimal solution, complexities in confirming network parameters as well as the slow learning rate. These drawbacks of the BP neural network were essentially eliminated through dimension reduction (Hinton and Salakhutdinov, 2006) and optimization algorithms for the local convergence and learning rate (Jin et al., 2000) to enhance data representation and optimize network parameters, respectively.

Presently, academics and practitioners highly prefer RNN and CNN deep learning techniques for FTS forecasting. RNN models use the input data for training and validation to forecast on FTS, while they come with the compromise of gradient vanishing and explosion during the training process (Hochreiter and Schmidhuber, 1997), resulting in loss of long-term information over time. CNN models are yet another deep-learning method that offers visual recognition capability and considerable reduction in parameter amount (Jiang et al., 2023) while they also come with the compromise of computational complexity in setting accurate kernel parameters, potentially leading to overfitting in the testing data. To address these issues, Hochreiter and Schmidhuber (1997) introduced LSTM neural networks that are capable of capturing long-term dependencies and alleviating the issue associated with gradient vanishing and explosion in RNNs. Likewise, the literature review of Sezer et al. (2020) emphasize that LSTM and its hybrid model variations outperform the other DL models due to its successful implementation of sequential data processing capabilities, which effectively capture long-term dependencies and temporal patterns in the FTS forecasting domain.

### **2.2.1. Long Short-Term Memory Neural Networks**

Hochreiter and Schmidhuber (1997) introduced Long Short-Term Memory (LSTM) networks as a specialized type of recurrent neural network (RNN) that is capable of remembering both short-term and long-term historical data. LSTM networks have been highly favored among deep-learning practitioners, especially for complicated tasks such as automatic speech recognition and handwritten character recognition. LSTM networks have a wide range of applications, including but not limited to language modeling and translation, speech recognition, sentiment analysis, predictive analytics, and financial time series analysis (Wu et al., 2016; Greff et al., 2016).

Due to their outstanding memory capabilities, LSTM networks particularly excel at managing time series data, making them ideal for financial time series (FTS) forecasting (Fang et al., 2023). The defining feature of LSTM networks is their memory cell architecture. These networks comprise LSTM units, which collectively form an LSTM layer. Each LSTM unit has cells with an input gate, an output gate, and a forget gate. These gates manage the flow of information, allowing the LSTM layer to evaluate the relevance of input data, forget irrelevant data, update the cell state, and provide appropriate outputs. This memory cell structure of LSTM networks allows each cell to keep critical values throughout different periods.

As LSTM networks are a specialized type of RNN, the optimization techniques and weight updates used are identical to those of RNN as well as the hyperparameters such as the number of layers, number of neurons per layer, learning rate, batch size, dropout rate, sequence length, and activation function (Reimers and Gurevych, 2017). Optimizing the hyperparameters of the LSTM networks can significantly increase the performance of the LSTM networks, for which the following optimization techniques are mainly employed: parameter sweeping, randomized search, and Bayesian Optimization techniques (Reimers and Gurevych, 2017).

Furthermore, principal component analysis (PCA) is a statistical technique that reduces the dimensionality of a dataset while keeping the majority of the variation. It determines the principal components, or the directions in which the data fluctuates the most, and converts the original variables into a new data set of uncorrelated variables sorted by the loadings of variance explained (Wen et al., 2020). With regards to LSTM networks, PCA increases performance by removing noise and redundant features of the raw data, lowering computational cost and increasing the model's capacity to detect important patterns in time-series data (Fang et al., 2020). This data preprocessing phase is especially useful in financial applications, which frequently use high-dimensional datasets and require computational efficiency.

Numerous studies in the literature (Sezer et al., 2020) have shown that LSTM networks are successful at predicting stock prices by leveraging their ability to represent temporal dependencies in FTS data. For instance, Chen et al. (2015) used LSTM networks with raw time series data to predict over 3000 Chinese stocks between 1990 and 2015 and found the accuracy of stock price prediction rose from 14.3% to 27.2% over other ML models. Likewise, Fischer and Krauss (2018) tested LSTM networks on S&P 500 index constituents between 1992 and 2015 for out-of-sample predictions and found that LSTM networks outperform the US stock market with an average daily return of 46 bps as well as other neural networks and logistic regression models in forecasting stock price changes. The study by Yan and Ouyang (2018) uses a hybrid model that combines LSTM networks with wavelet analysis to predict the daily closing price of the Shanghai Composite Index (SSE 180) and demonstrates the significant predictive power of LSTM networks on FTS. Furthermore, Siami-Namini et al. (2018) compared LSTM networks to ARIMA models and demonstrated that LSTM networks significantly outperformed ARIMA models, achieving an average reduction in error rates between 84% and 87% in forecasting stock prices.

### **2.3. Hypotheses**

Before performing the analyses of the thesis, I hypothesize that the monthly rebalanced equal-weighted (EW) long-short portfolios generated by the hybrid strategy of fundamental factor sorting and LSTM network will significantly outperform the S&P Europe 350 Index and the momentum factor. For the first, with the efficient screening of balance sheets for identifying and differentiating companies based on their sensitivity with the market movements, market capitalizations, valuation, profitability, growth,

investment, liquidity, and solvency metrics, as well as identifying price- and trend-related trading signals through LSTM networks, I anticipate to identify the best- and worst-performing constituents of the S&P Europe 350 Index accurately with accurate market-timings to outperform the passively managed the S&P Europe 350 Index through taking a long-short strategy. For the latter, the paper of Jiang et al. (2023) shows well-consistent findings for another neural network method (2D-CNN) outperforming the momentum factor. Aligned, as I will incorporate not only price or trend-related strategy but also balance sheet materials into this hybrid investment strategy, I expect to outperform the momentum factor by identifying and quickly adapting the strategy through monthly rebalancing to capitalize on potential short-term trend reversals, or market shifts.

Additionally, I expect that this hybrid strategy will outperform each factor of the 5-factor Fama-French model. The underlying reason of this hypothesis is similar to the hypothesis of outperforming momentum factor. I expect that by incorporating valid and stable balance sheet materials and technical indicators into the hybrid portfolio strategy can potentially allow outperforming relatively more restricted individual factors of the 5-factor model. Likewise, as this papers incorporates broader fundamental data pool relative to the 5-factor model along with 6 daily historical trading volume and price data variables, I hypothesize that the hybrid portfolio strategy through its holistic approach can outperform the 5-factor Fama-French model as a whole.

Therefore, in this thesis, I formulate the following four hypotheses:

Hypothesis 1: Monthly rebalanced EW long-short portfolios generated by the hybrid strategy of fundamental factor sorting and LSTM network will significantly outperform the S&P Europe 350 Index.

Hypothesis 2: Monthly rebalanced EW long-short portfolios generated by the hybrid strategy of fundamental factor sorting and LSTM network will significantly outperform the momentum factor based on the European equity market.

Hypothesis 3: Monthly rebalanced EW long-short portfolios generated by the hybrid strategy of fundamental factor sorting and LSTM network will significantly outperform each factor of the 5-factor Fama-French model based on the European equity market.

Hypothesis 4: Monthly rebalanced EW long-short portfolios generated by the hybrid strategy of fundamental factor sorting and LSTM network will significantly outperform the whole 5-factor Fama-French model based on the European equity market.

## CHAPTER 3 Data

In this thesis, I chose to investigate the historical constituents of the S&P Europe 350 Index from June 2019 to December 2023 where this period and these constituents are used for out-of-sample forecasting to implement the monthly rebalanced equal-weighted long-short hybrid portfolio strategy. Firstly, I gathered historical constituents of the index information from the Bloomberg Terminal from June 2019 to December 2023. Secondly, to identify and select which fundamental factors are valid and stable during the in-sample period to be used for the out-of-sample period forecasting, I gathered company-level 26 balance sheet materials (from a monthly to a quarterly basis) from the Thomson/Refinitiv Eikon database from December 2001 to November 2023. Thirdly, to train and validate technical indicators during the in-sample period through LSTM networks to do predictions for the out-of-sample period, I downloaded 6 security price-relevant daily technical indicators from Yahoo Finance from the 1st of January 2001 to the 29th of December 2023. Furthermore, I collected the monthly returns of securities from January 2001 to December 2023 from both Yahoo Finance and Thomson/Refinitiv Eikon databases to confirm that both the security price information and the monthly security returns were accurate. The monthly returns of securities are crucial for two reasons. Firstly, the monthly returns are used for the validation process of the balance sheet materials in which I check whether each of these monthly balance materials at time  $t$  is significantly correlated with the monthly returns at time  $t+1$  on a monthly basis on a security level during the in-sample period. Secondly, I later use the monthly returns of securities to assess the performance of the balance sheet-based sorted quintiles during the out-of-sample period.

Before the data-cleaning process, the raw sample dataset had 406 securities in total. To avoid the survivorship bias in the portfolio strategy, I included securities that were delisted, declared bankruptcy, or acquired between June 2019 to December 2023 (out-of-sample period). Likewise, to avoid look-ahead bias, I updated the investible universe after each rebalancing of the S&P Europe 350 Index itself; thus, I added or removed securities from the sample dataset once they became available or unavailable, respectively, as index constituents. In this regard, for the 26 balance sheet materials and the monthly returns of securities, the raw sample dataset has  $406 \times 27 \times 264$  (#securities, #variables, #months) = 2893968 company-level monthly observations. As for the 6 daily technical indicators, the raw sample dataset of the 406 securities includes  $406 \times 6 \times 4956,7$  (#securities, #variables, #average day availability) = 12074526 company-level daily observations. The differing day availability of stocks in the technical indicators dataset, particularly during the in-sample period, stems from the IPOs of the constituents occurring after the in-sample period began, signaling the importance of the data-cleaning process for the raw data as well as the importance of avoiding survivorship and look-ahead biases by updating the investible universe simultaneously with the index rebalancing periods.

### **3.1. Historical Constituents of the S&P Europe 350 Index**

S&P Europe 350 is an equity-based quarterly-rebalanced value-weighted index that tracks 350 industry-leading firms that are incorporated and listed in 16 different developed European markets. In these

markets, the securities are primarily chosen for this index based on their free-float market values and liquidity positions. More specifically, the securities are chosen if they have among the highest free-float market capitalization and the trailing 12-month average of share turnover. The latter is required to be at least 0.24 in each quarter for the current securities in the index, while it needs to be at least 0.30 for a potential constituent to be added to the index. Furthermore, S&P Global Ratings offers this index in two currencies: the Euro and the U.S. dollar. I decided to use the U.S. dollar-based version of the index in order to keep the currency consistent with the returns of the 5-factor of the Fama-French model and the momentum factor returns, where the rest are also calculated in the U.S. dollar.

The index was founded in 1998 (four years before the outset of the in-sample period) and is closely regulated by the European Union and the UK Financial Conduct Authority (FCA). The underlying reason for choosing the beginning of the in-sample period four years after the launch date of the index is to avoid the noise of the Dot-com bubble in the late 1990s and early 2000s. Particularly when training and validating the technical indicators by running LSTM neural networks, the immediate noise in the data could have posed the risk of inefficient performance of the LSTM networks to understand the patterns in the technical indicators. Aside from providing a diverse range of different European markets, the index also offers a broad range of industries (11 GICS sectors) which could allow flexibility in the hybrid portfolio strategy to capitalize on the mispricing of securities to a greater extent. Additionally, the consistent number of 350 constituents is another reason I choose this index over similar indices as the hybrid portfolio strategy would achieve ideal diversification of 12 stocks monthly for portfolio creation after fundamental factor sorting and running the LSTM neural network. The historical constituents of the index are limited to the out-of-sample period of June 2019 to December 2023. For this period, I downloaded the historical constituents of the index in 19 quarters from Bloomberg Terminal and I merged these quarterly historical constituents data and investigated the quarters of each security being included and excluded from the index. In this part, it is crucial to set the default of all existing constituents in June 2019, regardless of whether they were added to the index earlier or in the quarter of June, as they were included in June 2019. After identifying and setting the first and last (if removed from the index) observation quarters of historical constituents in the index, I created a binary variable to indicate whether the given security is present or not in the investable universe for the hybrid portfolio strategy, 1 indicating presence. This step is essential to prevent survivorship and look-ahead biases. Furthermore, the underlying reason for choosing this out-of-sample period of June 2019 to December 2023 for the historical constituents was due 80/20 rule. This rule is widely used in deep learning methods to improve the accuracy of the models and it suggests that the training and validating period accounts for 80% of the total period whereas the out-of-sample testing period only composes 20% of the total period.

### **3.2. Fundamental Data**

The first stage of formulating the hybrid portfolio strategy is to sort the historical constituents based on their balance sheet materials and construct monthly rebalanced equal-weighted quintile portfolios. To extensively distinguish and assess the companies based on their financial stability, it is crucial to collect diverse fundamental elements that can considerably represent the companies' financial health in that respective month. In this regard, after consolidating the historical constituents list of the index during the out-of-sample period, I scrutinized the Thomson/Refinitiv Eikon database for data availability of balance sheet materials for the given 406 constituents from December 2001 to December 2023. I checked whether each potential balance sheet material has at most 50% of the total observations missing; thus, each balance sheet material needs to have at least  $406 * 264 * 0,5$  (#securities, #months, #maximum 50% of the data missing) = 53592 monthly observations. As a result, I ended up gathering 26 company-level balance sheet materials from gross values to multiples where all are relevant for the fundamental analysis of the constituents as shown in Table 1. Each of these monthly balance sheet materials will first be tested for their correlation with the next month's monthly returns on a security level during the in-sample period. The ones that pass this validity test will be further checked for their stability during the in-sample period. Furthermore, I will check the correlation matrix of valid and stable balance sheet materials. In case there is either a high correlation among balance sheet materials or a high number of balance sheet materials passing both tests (validity and stability), I will use principal component analysis to reduce balance sheet material dimensions and avoid potential multicollinearity. Therefore, when I gathered these 26 balance sheet materials in Table 1, I did not immediately eliminate any for their potential correlation with other balance sheet materials. As displayed in Table 1, 26 balance sheet materials are categorized as follows: 9 on valuations, 6 on profitability, 5 on growth and investment, 4 on liquidity and solvency, 1 on volatility, and 1 on size.

**Table 1.** Fundamental & technical data categorization.

| Balance Sheet Materials (26)  |   |   |   |                | Technical (6)   |  |
|---|---|---|---|----------------|-----------------|--|
| Valuation (9)   | Profitability (6)   | Growth & Investment (5)   | Liquidity & Solvency (4)                            | Volatility (1) | Size (1)        | Price and Volume (6)   |
| M/B,<br>P/E,<br>P/B,<br>P/C,<br>EV<br>/EBITDA,<br>EV/EBIT,<br>EV/SALES,<br>Dividend<br>Yield,<br>Total Assets | Operating Profit<br>Margin,<br>Gross Profit<br>Margin,<br>Net Margin,<br>ROE,<br>ROIC,<br>ROA | 12-Month<br>Forward P/E,<br>BV per Share<br>Growth,<br>Net Income<br>Growth,<br>Net Margin<br>Growth,<br>$\Delta$ Investments | Current Ratio,<br>Quick Ratio,<br>D/E,<br>Cap Ratio | Beta           | Market<br>Value | Open,<br>High,<br>Low,<br>Close,<br>Adjusted<br>Close,<br>Volume |

Table 1 illustrates the balance sheet materials used and their categorization from public corporate filings of the historical constituents of the S&P Europe 350 Index retrieved from the Thomson/Refinitiv Eikon database. M/B is the market-to-book ratio; P/E is the price-to-earnings ratio; P/B is the price-to-book ratio; P/C is the price-to-cash ratio; EV is the enterprise value; ROE is the return on equity; ROIC is the return on invested capital; ROA is the return on assets; D/E is the debt-to-equity ratio; BV is the book value. The detailed descriptions of the balance sheet materials can be accessed from Thomson/Refinitiv Eikon database.

### 3.3. Technical Data

The second stage of formulating the hybrid portfolio strategy is to train, validate, and forecast price movements through trading volume and price data by LSTM neural networks for the securities in the bottom and top quintiles in the first stage using fundamental data to construct a monthly rebalanced equal-weight high-low portfolio. To improve the performance of LSTM neural networks for the training, validation, and prediction processes, it is crucial not to limit the price information to only one metric, such as daily closing price, to allow LSTM neural networks to better understand long-term patterns and short-term temporal dependencies from various sources. For this purpose, given the historical constituents list of the index during the out-of-sample period, I downloaded daily opening, high, low, closing, adjusting closing prices, and volume (OHLCAV) data from Yahoo Finance for the given 406 constituents from the 1st of January 2002 to the 29th of December 2023 as shown in Table 1. Daily opening and closing prices indicate the price level of the first and last trade, respectively, of the security during an open-market day. Daily high and low prices show the price level of the highest and lowest trade, respectively, of the security during the day. The daily adjusted closing price is a retroactive measure of the daily closing price that takes corporate decisions, such as dividend distributions, stock

splits, and stock issuances, into consideration. The underlying reason for choosing Yahoo Finance for collecting historical trading data is its accessibility through the `yfinance` library, an open-source Python library, and its reliability. For the latter, I crosschecked the end-of-month daily closing price downloaded on a security level for each constituent from the `yfinance` library in Python with the values collected from the Thomson/Refinitiv Eikon database and concluded that the data was majorly same with minor differences for which I corrected manually. Furthermore, Yahoo Finance provides adjusted closing price historical data which allows to account for corporate actions, such as dividend distribution and stock splits, that can affect the price of a security directly. Through this, the historical price data assures to incorporate the accurate reflection of market and corporate conditions on a security level.

I automated the data collection process using the `yfinance` and `pandas` library in Python to download the daily OHLCAV data for the given 406 historical constituents from the 1st of January 2002 to the 29th of December 2023. In this regard, I first created an exchange mapping for the ticker suffixes between Yahoo Finance and Bloomberg Terminal due to differences in security tickers between the two platforms. After making the exchange mapping list and changing the stock suffixes for Yahoo Finance, there needs to be a manual crosscheck for those stocks that the ticker itself changes between the two platforms. Followingly, I used a custom function of `ohlcv.data` which contains an empty list to hold all the data frames for tickers in one list. This is a critical step to organize the OHLCAV data of tickers into a panel data format, preparing for the application of the LSTM neural networks and PCA easily. With the new list of tickers specifically for Yahoo Finance, I downloaded the OHLCAV data for each ticker using a loop function and concentrated each ticker data frame into one data frame through the `pd.concat` function. Besides, at the end, I use the `reset_index` function to transform the daily date information into a column. This is also a crucial step to ensure each observation in the panel data is referenced by both the respective ticker and date information. Moreover, separating and saving the panel data of daily OHLCAV into several chunks, in 3 different Excel sheets, could be required as the daily OHLCV data has over 12 million company-level daily observations, which could potentially exceed the file boundaries.

### **3.4. External Data**

Hypotheses from 1 to 4 challenge the performance of the hybrid portfolio strategy to outperform the S&P Europe 350 Index, the momentum factor, and factors of the 5-factor Fama-French model on a monthly basis. To this end, I collected the end-of-month prices of the S&P Europe 350 Index from the FactSet database from the 29th of March 2019 to the 29th of December 2023 to calculate the monthly returns of the index from June 2019 to December 2023. Ensuring the given index prices are in the U.S. dollar is crucial to keep the currency consistent. As for the momentum factor and factors of the 5-factor Fama-French model, I gathered the monthly returns of these factors for 16 European markets from the Kenneth R. French data library. All the monthly returns are provided in U.S. dollars.

While choosing the Kenneth R. French data library, I ensured that the 16 European markets used in this data library are reflective of the investible universe of the hybrid portfolio strategy which is based on the S&P Europe 350 Index, which also uses securities from 16 developed European markets. Both the index and Kenneth R. French data library share 15 European markets in common. The only difference stems from Luxembourg and Greece in which the S&P Europe 350 Index uses Luxembourg while the Kenneth R. French data library uses the Greece market instead. This discrepancy would not disrupt the findings as Luxembourg only accounts for 0.1%, in terms of market capitalization, of the S&P Europe 350 Index in the country breakdowns. Likewise, in the value-weighted portfolio of the 5-factor Fama-French model, the stock exchange of Greece, in terms of market capitalization, is the lowest among the 16 European markets.

### **3.5. Data Cleaning Process**

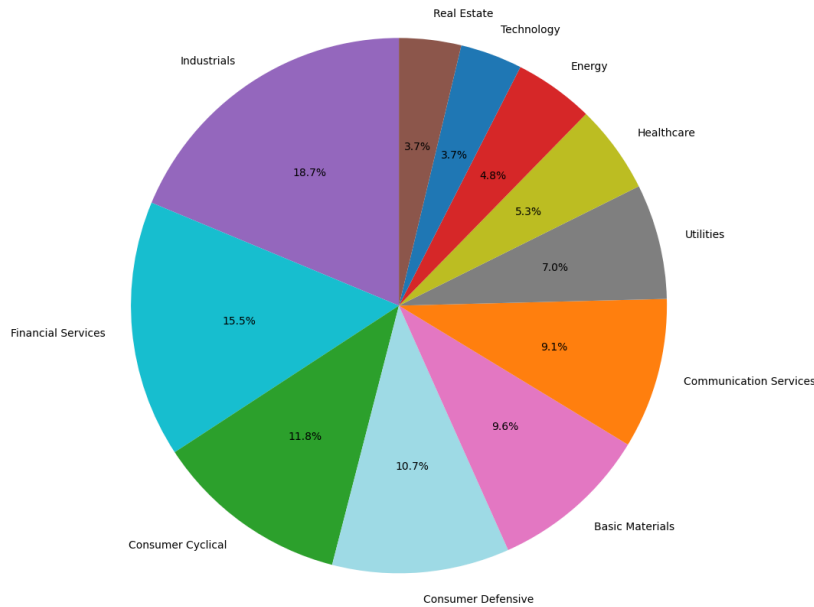
After the technical data collection of daily OHLCAV data, there is no missing data observed; thus, there are 12074526 company-level daily observations downloaded from Yahoo Finance through the yfinance library of Python for the 406 historical constituents. However, fundamental data shows a considerable amount of missing data over 20 years. Taking the IPO dates into consideration for both in-sample and out-of-sample periods, securities, that do not have any fundamental data or have at least 50% of the fundamental data missing during the out-of-sample period, were removed from the sample dataset due to their data unavailability. As a result, there are 340 securities left in the data sample. After also dropping the missing observations, the sample dataset of 340 securities for the fundamental data has 2110179 company-level monthly observations for 27 variables (including the monthly returns of securities). Likewise, the technical data, as a result, has 10169040 company-level daily observations for 6 trading volume and price variables.

The underlying reasoning for removing such securities is that if the security does not steadily provide fundamental data information during the out-of-sample period on a monthly basis, adding and removing that security to the investible universe for the monthly balance sheet-based quintile would not reflect a reliable and consistent performance of the hybrid portfolio strategy. Secondly, in case there is a high correlation between balance sheet materials and/or a high number of balance sheet materials passing both validity and stability tests in the findings, principal component analysis (PCA), accompanied by the pre-processing step of standardization of the balance sheet materials, is employed, which does not allow for unbalanced panel data. Therefore, I decided to drop securities that do not have any fundamental data or have at least 50% of the fundamental data missing during the out-of-sample period. To this end, I also dropped the same securities from the technical dataset that was shown missing in the fundamental data. Furthermore, to eliminate the outliers, which could potentially create bias or noise in our findings, I censored all fundamental variables using the 1st and 99th percentiles on a security level. Lastly, the technical data containing the OHLCAV data is standardized before the PCA analysis and normalized before the LSTM neural network. The underlying motivations of standardization as the

pre-processing step of PCA analysis are for trading volume and price data variables to contribute equally to create a price index and avoid variance across 6 different data variables, for which PCA analysis is sensitive. Likewise, the normalization step is crucial to improve the convergence rate of the LSTM neural network, in which the training and validation results of the LSTM model are steadily improving after each epoch. Lastly, while creating a balance sheet score index, in order not to over or undervalue any security – due to potentially missing several variables on a specific month – I decide to use median (mean) imputation *on a security level* if the distribution of the respective variable is skewed (normal). After executing skewness and kurtosis tests for normality, I identified that all fundamental data variables have skewed distribution; thus, I used the median imputation on a security level on the clean dataset of 340 securities.

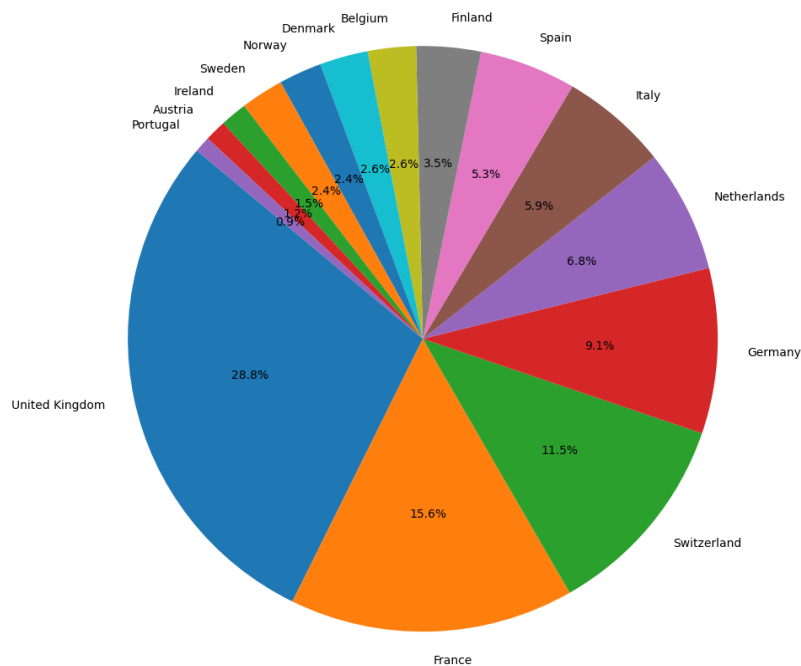
### **3.6. Summary Statistics**

Before the methodology of the paper, it is crucial to provide summary statistics of the clean sample dataset of the fundamental and technical data to better understand the nature of the investible universe, data variables, and their correlations. In this regard, the clean data sample is composed of 340 securities in total over the period from June 2019 to December 2023, accounting for the quarterly rebalancing of the index. Figure 1 illustrates the sectoral breakdown in terms of GICS classification, and the clean dataset, most notably, includes all 11 sectors. The sectoral diversity of the dataset offers an opportunity for the hybrid portfolio strategy to capitalize on any sectoral opportunities. 18.7% of the dataset falls into the industrial sector as the largest sector as shown in Figure 1. The industrials is a broad sector including but not limited to securities that mainly operate in the production and distribution of capital goods, transportation, and logistics. Likewise, financial services, consumer cyclical (discretionary), and consumer defensive (staples) considerably account for 15.5%, 11.8%, and 10.7% of the dataset, respectively. Having fairly close percentages of consumer cyclical and defensive sectors allows hedging for economic cycle conditions, in which the cyclical (defensive) sector includes securities that are (un)correlated with the market conditions. Furthermore, the basic materials and communication services sectors have a moderate representation of 9.6% and 9.1%, respectively, in the overall dataset. Lastly, from utilities to real estate sectors, the other 5 sectors only account for 24.5% of the total dataset as illustrated in Figure 1.



**Figure 1.** Pie chart of the sector breakdown of the sample dataset.

As aforementioned, the S&P Europe 350 Index includes securities from 16 developed European markets (the UK included). Thus, it's crucial to check whether the clean dataset represents the geographical diversity similar to that of the index. In this regard, Figure 2 illustrates the pie chart of the country breakdown of the dataset, which includes 15 of the 16 developed markets. The only country missing in the clean dataset is Luxembourg. However, this would not hamper the geographical diversity of the dataset as Luxembourg only accounts for 0.1% of the index itself in terms of country breakdown. Furthermore, when benchmarking the performance of the hybrid portfolio strategy against the 5-factor Fama-French model, the issue of Luxembourg missing from the 5-factor model, as a result, is resolved. As illustrated in Figure 2, 28.8% of the securities in the dataset are incorporated and listed in the United Kingdom. Followingly, France, Switzerland, and Germany account for 15.6%, 11.5%, and 9.1% of the dataset, respectively. Other countries, such as the Netherlands, Italy, and Spain hold relatively smaller yet still significant portions ranging from 6.8% to 5.3% of the dataset. The remaining 8 countries make up 17.1% of the dataset, each contributing at most or less than 3.5%.



**Figure 2.** Pie chart of the country breakdown of the sample dataset.

In addition to the security characteristics of the dataset in terms of sector and country distribution, it is also crucial to provide descriptive statistics of the fundamental and technical data to analyze the extent of variability and distributional attributes among securities in the dataset that can influence investment decisions and risks associated with them. In Table 2, the first 27 rows, from market-to-book ratio to adjusted monthly return, provide descriptive statistics on the fundamental data, while the last 6 rows, from opening price to volume, provide those on the technical data. The dataset variables, particularly fundamental ones, show a significant level of security heterogeneity with high standard deviation levels. This can be explained by the sectoral and geographical differences among securities.

In Table 2, the market-to-book ratio, a valuation metric indicating market valuation relative to the book value of a security, has a mean and standard deviation of 3.05 and 3.14, respectively, as well as minimum and maximum values of 0.15 and 20.48. Likewise, in other valuation metrics such as price-to-earnings, price-to-cash, dividend yield, and enterprise value multiples, there is a similar wide distribution across securities in the dataset, implying the existence of growth and value stocks in the dataset. Furthermore, the market value, a size metric indicating the total market capitalization of security, has a mean and standard deviation of approximately \$25.2 and \$40.8 billion, respectively, as well as minimum and maximum values of approximately \$0.3 and \$246.2 billion, respectively. This highlights the extensive size range of securities in the dataset from small-caps to mega-caps securities. Likewise, profitability metrics such as operating profit margin, ROE, ROIC, ROA, and net margin, have positive means as well as standard deviations that are either similar or higher levels compared to their respective means. This high standard deviation, accompanied by negative minimum values in these profitability metrics, shows the existence of robust and weak profitable securities in the dataset.

Similarly, growth metrics such as the 12-month forward P/E, net income growth, and net margin growth indicate overall optimistic consensus in future earnings growth of securities in the dataset with positive averages, yet their respective substantial standard deviation and tails reflect uncertainty and volatility across different sectors and markets. As for the market behavior of securities in the dataset, the beta, a volatility metric indicating the systematic risk of the respective security, shows a mean of 1.02, which indicates the dataset is slightly riskier compared to the overall European market. In addition to the standard deviation of 0.5, the minimum (0.05) and maximum values (2.60) of the beta present the existence of low and high-risk securities in the dataset. Lastly, the daily trading volume of securities reveals substantial swings of securities in the dataset ranging from as low as 0 to over 137 million. The significant gap between the 25th and 75th across the majority of variables reassures the security heterogeneity of the dataset. Furthermore, Table 1A displays the p-values for the skewness and kurtosis for each respective fundamental and technical data variable where every variable except the monthly returns of securities exhibits statistically significant skewness at a 1% level.

**Table 2.** Descriptive statistics of the fundamental & technical data.

| Variable          | N     | Mean  | SD    | Min    | 25th  | Median | 75th  | Max     |
|-------------------|-------|-------|-------|--------|-------|--------|-------|---------|
| Market-to-Book    | 77892 | 3.05  | 3.14  | 0.15   | 1.21  | 2.10   | 3.64  | 20.48   |
| Market Value      | 80291 | 25.23 | 40.79 | 0.27   | 4.70  | 10.01  | 26.33 | 24.62   |
| Price-to-Earnings | 73543 | 22.57 | 24.22 | 3.50   | 11.60 | 16.70  | 24.60 | 193.56  |
| Dividend Yield    | 80290 | 2.92  | 2.19  | 0.00   | 1.38  | 2.57   | 4.12  | 10.68   |
| Operating Profit  | 83517 | 14.14 | 14.18 | -37.75 | 6.12  | 12.20  | 19.73 | 64.58   |
| Margin            |       |       |       |        |       |        |       |         |
| Beta              | 79953 | 1.02  | 0.50  | 0.05   | 0.67  | 0.97   | 1.33  | 2.60    |
| Total Assets      | 83685 | 11.42 | 31.52 | 0.16   | 0.53  | 1.55   | 55.78 | 207.060 |
| ROE               | 81817 | 15.49 | 19.52 | -44.83 | 7.11  | 13.65  | 21.69 | 108.81  |
| ROIC              | 82653 | 9.99  | 11.17 | -24.56 | 4.18  | 8.39   | 13.97 | 57.67   |
| ROA               | 81942 | 6.21  | 6.94  | -15.15 | 1.92  | 5.34   | 9.17  | 33.96   |
| Net Margin        | 83541 | 9.99  | 21.62 | -80.95 | 3.22  | 7.57   | 13.95 | 138.69  |
| Current Ratio     | 66669 | 1.52  | 0.95  | 0.41   | 0.97  | 1.29   | 1.74  | 6.51    |
| Quick Ratio       | 66741 | 1.05  | 0.67  | 0.21   | 0.67  | 0.89   | 1.22  | 4.70    |
| Debt-to-Equity    | 83625 | 1.38  | 2.03  | -1.28  | 0.36  | 0.69   | 1.45  | 11.80   |
| Price-to-Book     | 78573 | 3.04  | 3.22  | 0.20   | 1.20  | 2.07   | 3.57  | 21.33   |
| Price-to-Cash     | 78573 | 11.16 | 15.00 | -40.13 | 5.23  | 9.10   | 14.43 | 106.07  |
| EV/EBITDA         | 74127 | 11.58 | 10.06 | -0.76  | 6.39  | 8.98   | 13.01 | 72.53   |
| EV/EBIT           | 77978 | 14.92 | 10.98 | -10.97 | 9.35  | 12.41  | 17.17 | 72.04   |
| EV/SALES          | 78397 | 3.71  | 4.97  | 0.19   | 0.89  | 1.85   | 3.92  | 26.73   |
| Gross Profit      | 70046 | 38.37 | 22.67 | 3.08   | 20.08 | 33.79  | 53.83 | 97.11   |
| Margin            |       |       |       |        |       |        |       |         |

|                     |         |         |          |        |        |         |         |           |
|---------------------|---------|---------|----------|--------|--------|---------|---------|-----------|
| 12-Month Forward PE | 79356   | 15.39   | 8.74     | -13.21 | 10.48  | 13.94   | 18.58   | 56.67     |
| BV per Share        | 81385   | 8.79    | 27.07    | -54.96 | -1.91  | 5.97    | 14.84   | 175.27    |
| Growth              |         |         |          |        |        |         |         |           |
| Net Income          | 69204   | 42.36   | 160.05   | -91.19 | -12.24 | 10.42   | 38.44   | 1192.71   |
| Growth              |         |         |          |        |        |         |         |           |
| Net Margin          | 69097   | 31.12   | 148.83   | -92.28 | -16.69 | 2.46    | 24.73   | 1126.40   |
| Growth              |         |         |          |        |        |         |         |           |
| Cap Ratio           | 83649   | 0.43    | 0.24     | 0.00   | 0.26   | 0.40    | 0.58    | 1.04      |
| ΔInvestments        | 83344   | 0.12    | 0.93     | -1.52  | 0.00   | 0.00    | 0.00    | 7.83      |
| Monthly Return      | 80291   | 0.01    | 0.08     | -0.23  | -0.04  | 0.01    | 0.05    | 0.25      |
| Open                | 1694840 | 309.96  | 466.69   | 1.08   | 21.46  | 66.94   | 368.18  | 1448.22   |
| High                | 1694840 | 312.51  | 469.37   | 1.09   | 21.74  | 67.66   | 373.00  | 1452.92   |
| Low                 | 1694840 | 306.71  | 462.57   | 1.08   | 21.16  | 66.12   | 362.84  | 1437.66   |
| Close               | 1694840 | 309.80  | 466.40   | 1.09   | 21.46  | 66.90   | 368.03  | 1447.05   |
| Adjusted Close      | 1694840 | 246.81  | 407.40   | 0.64   | 14.76  | 48.77   | 246.11  | 1396.39   |
| Volume              | 1694840 | 5899264 | 14832580 | 0.00   | 426228 | 1359730 | 4441506 | 137374500 |

Table 2 provides the market value, total assets, and change in investment in billions. Other fundamental data is provided in ratios, excluding margins and growth metrics which are provided in percentages.

Furthermore, it is crucial to provide and analyze the correlation matrix of the fundamental data variables for two reasons: detecting potential multicollinearities and employing PCA analysis. The first is important as 24 of 26 fundamental variables share a similar balance sheet metric and are in the same balance sheet category such as valuation, profitability, growth & investment, and liquidity & solvency as shown in Table 1. The latter is a crucial optional step in case there is a high number of fundamental data variables passing both tests of validity and stability. In Figure 1A, the high correlation of 0.97 between market-to-book and price-to-book is reasonable as they are the same variables. I included both as they are from different data sources in the database of Thomson/Refinitiv Eikon and their standard deviations and tail values are not the same, which could lead to the elimination of one, or both, variables during the validity and stability tests. Likewise, as assumed, profitability metrics (ROE, ROIC, ROA, other margins), liquidity metrics (current and quick ratio), solvency metrics (debt-to-equity and capitalization ratio), and valuation metrics (EV multiples) have high correlation levels in their own category as illustrated in the correlation matrix of Figure 1A.

## CHAPTER 4 Method

The objective of devising such a hybrid portfolio strategy that integrates principal component analysis both for fundamental factor sorting (optionally) and for Long Short-Term Memory (LSTM) neural networks for the technical data is to pursue a consistent abnormal return-generating investment strategy on a risk-adjusted basis. Through this novel hybrid portfolio strategy, I aim to optimize the portfolio performance by relying not only on fundamental analysis but also on technical analysis and identifying the advantages and shortcomings of different analyses. The strength of the methodology employed in this study stems from two different strategies (fundamental factor sorting and the LSTM model for technical indicators) compensating for each other's limitations. The LSTM neural network is subject to lower predictive accuracy when markets are highly volatile, and this study addresses this limitation by incorporating fundamental factor sorting on a company level and creating a long-short portfolio that promises lower overall portfolio volatility (Blackrock, 2024). On the other hand, the fundamental analysis is also subject to not reflecting the fair value of equity, particularly in the short-term, due to mispricing or speculations that take place in the equity markets. Thus, incorporating technical analysis through trading volume and price data and the LSTM neural network, can inform on the price trends and swings and address this mispricing issue to increase the performance of fundamental analysis. Furthermore, this study uses a relatively narrow investible universe of the S&P Europe 350 Index constituents, which consists of 350 industry-leading firms with the highest market capitalization and market liquidity levels. Consequently, using this investible universe and integrating fundamental and technical analyses through PCA and LSTM neural networks, this study aims to outperform relevant market benchmarks, such as the index itself, Fama-French 5-factor returns, and momentum factor in terms of risk-adjusted returns.

The dataset is divided into two timeframes of in-sample and out-of-sample periods, in which the first is used for training and validation and the latter is used for the hybrid portfolio construction and its performance evaluation. When allocating the timeframes of in-sample and out-of-sample periods, I used the 80/20 train-test split ratio, particularly to optimize the learning rate and performance of the LSTM neural network. The 80/20 train-test split is a common practice in machine learning models to train and validate the dataset to have increased performance, and the finding of Gholamy et al. (2018) provides statistical proof of this 80/20 split optimizing the training and test periods as well as minimizing overfitting of the dataset. The in-sample refers to the period in which the fundamental data variables are screened for their respective validity and stability for the creation of the balance sheet score index (on a company level, monthly basis), whereas the technical data is used for training and validating processes of the LSTM neural network. The out-of-sample refers to the period in which the balance sheet score indexes are sorted to create monthly rebalanced quintile equal-weighted (EW) portfolios, whereas the LSTM neural network model does predictions of the price index (on a company level, daily basis). Subsequently, the outputs of the monthly rebalanced quintile EW portfolios and LSTM predictions are

combined to construct a monthly rebalanced hybrid long-short EW portfolio. The in-sample period for the fundamental dataset is from December 2001 to April 2019, while the in-sample period for the technical dataset is from January 2002 to May 2019. The out-of-sample period for the fundamental dataset is from May 2019 to November 2023, whereas the out-of-sample period for the technical dataset is from June 2019 to December 2023.

To avoid survivorship and look-ahead biases, the investible universe is quarterly updated in March, June, September, and December based on the rebalancing frequency of the index. Thus, to deal with survivorship bias, I include securities in the dataset that were delisted, declared bankruptcy, or acquired during the out-of-sample period. Likewise, to deal with look-ahead bias, I added or removed securities from the sample dataset once they became available or unavailable, respectively, as constituents of the S&P Europe 350 Index during the out-of-sample period. Survivorship bias emerges when the dataset only accounts for securities that have maintained their operations during the out-of-sample period, which is 4.5 years in this study. Not addressing survivorship bias could significantly overestimate the performance of the hybrid portfolio strategy and hamper the short-leg returns of the portfolio. Besides, look-ahead bias emerges when the hybrid portfolio strategy uses future information, that is already out, as if it was available at the time such as the inclusion or exclusion of security from the investible universe beforehand or afterward, respectively, before the index rebalancing. Thus, to assess the performance of the hybrid portfolio strategy fairly, it's critical to address these biases.

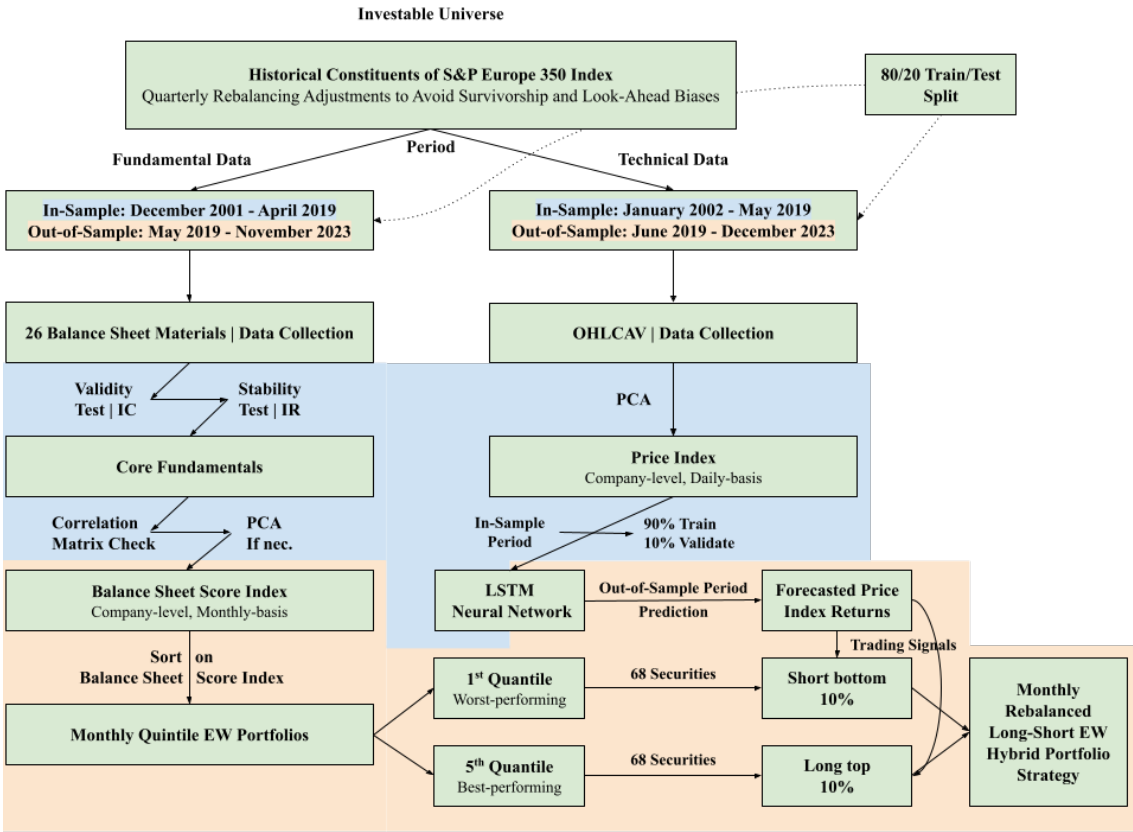


Figure 3. Diagram of methodological framework of the hybrid portfolio strategy.

As illustrated in Figure 3, the research methodology of this thesis is constructed around PCA and the LSTM neural network to formulate the hybrid portfolio strategy. In the fundamental factor sorting section, the initially collected 26 balance sheet materials are screened based on their validity and stability for which information coefficient (IC) and information ratio (IR) are used, respectively, during the in-sample period (Zhang et al., 2021). IC is given by the correlation of each fundamental data variable with the next month's return on a company level. Thus, it is essentially used to identify whether a monthly fundamental data variable has predictive power for the next month's returns on a company level (Margenot et al., 2024). In case the absolute value of IC of a variable is higher than, or equal to, 0.01, then it is considered as a candidate valid fundamental data variable. If not, the respective variable is dropped from the balance sheet material pool for the creation of the balance sheet score index and fundamental factor sorting process. The absolute value of IC allows me to take fundamental data variable(s) having a significant negative relationship with the next month's returns into account, which would provide a more objective distribution for the balance sheet score index. Thus, IC is calculated as follows:

$$IC_{ft} = correlation(e_{ft}, r_{f(t+1)}) \quad (1)$$

in which  $IC_{ft}$  stands for the information coefficient of variable  $f$  at time  $t$ ;  $e_{ft}$  stands for the monthly exposure to the variable  $f$  at time  $t$  on a security level;  $r_{f(t+1)}$  stands for the monthly return of the security at time  $t + 1$  on a security level. The benchmark of 0.01 for IC values could seem conservative at first; however, there are two underlying reasons for setting a lower benchmark. Firstly, the t-statistics of the null hypotheses for an IC value of 0.01 is statistically significant from 0 at a 1% level. Secondly, given the breadth of balance sheet materials, incorporating various financial metrics – with a relatively lower IC value if only stable - could reflect the financial health of a company better on the balance sheet score index.

Followingly, only for those valid fundamental variables with absolute values of IC higher than, or equal to, 0.01 during the in-sample period, I also check for their respective IC stability through information ratio (IR) over the in-sample period. This is a crucial step to ensure that the valid fundamental variables consistently have a statistically significant correlation for the next month's returns over the in-sample period for the out-of-sample period. While IR is conventionally used as a risk-adjusted performance ratio for active investing strategies, IR, in this case, simply provides the risk-adjusted value of IC to check for the stability of each fundamental variable. I use the IR formula of Qian and Hua (2004), who define IR as the mean of IC over the standard deviation of IC. As IC values for fundamental variables are on a monthly frequency and so are their standard deviations, I chose the stability benchmark of 0.289 (equal to an annualized IR of 1) for IR values for each respective variable. Any valid variable that has an absolute value of IR higher than, or equal to, 0.289 is considered to be a stable variable. Thus, IR is calculated as follows:

$$IR_f = \frac{\text{average } IC_f}{\text{std } IC_f} \quad (2)$$

in which  $IR_f$  stands for the information ratio of variable  $f$ ; average  $IC_f$  stands for the average of the information coefficient of variable  $f$  (Equation 1);  $\text{std } IC_f$  stands for the standard deviation of the information coefficient of variable  $f$  (Equation 1). Intuitively, a greater absolute value of  $IR_f$  indicates a variable is more stable in terms of its correlation with the next month's return. The fundamental variables that have at least an absolute  $IC$  value of 0.01 and at least an absolute  $IR$  value of 0.289 are considered to be valid and stable variables (core fundamentals). These core fundamentals are, therefore, selected to be used for the calculation of the balance sheet score index.

For those fundamental data variables passing both  $IC$  and  $IR$  tests, I check the correlation matrix (Figure 3) to identify whether any correlation appears among those variables to address multicollinearity concerns. In case there is either a high correlation level(s) among core fundamental variables or a high number of core fundamentals, I will employ principal component analysis (PCA) to prevent multicollinearity and reduce dimensions to create a balance sheet score index on a company level on a monthly basis for the out-of-sample period. Subsequently, I create monthly equal-weighted quintiles sorted on these balance sheet score indexes, in which the first quantile  $Q1$  (fifth quantile  $Q5$ ) represents the worst-performing (best-forming) securities, for the out-of-sample period.

As for the technical data section, the initial 6 historical trading volume and price data variables are processed through principal component analysis (PCA) to reduce dimensions in order to create a daily price index on a company level for the in-sample period. The underlying reason for using PCA to reduce dimension and followingly creating one variable price index that is again on a daily basis on a company level is to lower the computational efforts when running the LSTM model. Training and validating above 10 million data points is both computationally intensive and inefficient. Instead, deriving a price index through PCA that would capture a high degree of variation of the prominent features of the OHLCAV data would increase the performance of the LSTM model by yielding out-of-sample predictions in a timely and computationally effective manner. Besides, the OHLCAV data variables are inherently highly correlated, so reducing dimensions through PCA would be an optimal method. After creating such a daily price index, the LSTM neural network is employed to train and validate the daily price indexes for the in-sample period to provide out-of-sample daily price index forecasts. The initial 90% of the in-sample period is allocated for the training process of the LSTM, and the remaining 10% of the in-sample period is allocated for the validation process of the LSTM. The training process refers to the LSTM model using the initial 90% in-sample daily price index data to learn long-term patterns in the price index and to adjust the model's weights by using the Adam optimizer that minimizes mean squared error (MSE) loss. The validation process refers to the LSTM model using the remaining in-sample daily price index data to simultaneously assess the training performance of the model through testing the model's performance on the price index data that wasn't fed for the training process. Thus, the validation process is critical to avoid overfitting the price index data. After the training

and validation process for the in-sample period, the LSTM model provides the out-of-sample prediction of the daily price index data, from which I calculate the monthly returns of securities.

Given the monthly equal-weighted quintiles sorted on balance sheet score indexes and forecasted daily price index data through the LSTM model, I take a long position in the top 10% of Q5 securities, while I take a short position in the bottom 10% of Q1 securities, both on a monthly basis, to construct monthly rebalanced long-short equal-weighted portfolio in the out-of-sample period.

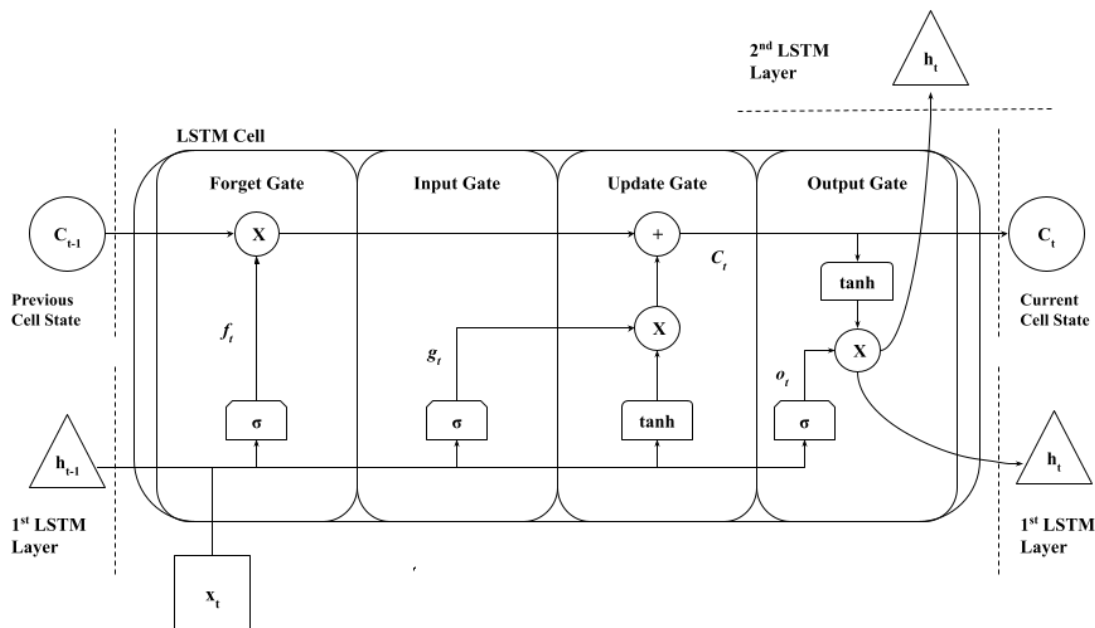
#### **4.1. Principal Component Analysis (PCA)**

PCA is a common statistical technique to reduce dimensions while maximizing the variance of the original data through principal components. Given the nature of (highly) correlated variables in the technical and fundamental data, employing PCA becomes critical to avoid multicollinearity, reduce noise in the data, increase the computational performance of the LSTM model, and reduce dimension to create daily price index and monthly balance sheet score index on a security level. Through the PCA technique, core fundamentals and OHLCAV data are both reduced into a smaller set of principal components which have a covariance of 0. For this aim, the following approaches were taken for both core fundamentals and OHLCAV data. While technical panel data had no missing data, core fundamentals panel data was unbalanced and had limited missing data after the data cleaning process, for which each core fundamental variable has median imputation on a security level. The median imputation is preferred over the mean imputation as all fundamental variables show skewed distributions (Table 1A). The imputation is required as PCA is sensitive to unbalanced panel data. Following that, each core fundamental variable and OHLCAV data is standardized as a pre-processing step to ensure that relatively large-scale variables do not distort the results of PCA. Subsequently, I run PCA analysis on STATA for both fundamentals and technical data separately. To decide on the number of principal components to be used, I used the Kaiser Rule (i.e. eigenvalue of 1), which dictates that each component at least comprise as much information as one respective variable. Eigenvalues simply inform on the magnitude of variance contained within one principal component; thus, any component with an eigenvalue less than 1 does not contribute any additional value relative to the variable and indicates that it is redundant. After determining the number of principal components for each core fundamental and OHLCAV data, I create new variables corresponding to these principal components with their given eigenvectors on both core fundamental and OHLCAV panel data. The eigenvectors are simply loadings of each variable on principal components. After the creation of principal component variables in both panel data, I check the proportion of variance explained of each principal component in which these proportions represent the weights of each principal component when generating the monthly balance sheet score index for core fundamentals and daily price index for OHLCAV data. Thus, the monthly balance sheet score index incorporates all valid and stable fundamental variables while avoiding multicollinearity and noise in the data. Likewise, the daily price index incorporates all 6 daily trading

volume and price variables while avoiding multicollinearity and lowering the computational efforts for the LSTM model.

#### 4.2. LSTM Neural Network

The LSTM neural network is particularly optimal when dealing with financial time-series (FTS) data forecasting due to its capability of not only learning both long- and short-term patterns but also retaining this information over long-term periods. In other words, the LSTM model incorporates both recent price movements and its long-term captured understanding of price movements and trends when doing out-of-sample FTS forecasting. Thus, LSTM overcomes the shortcomings of Recurrent Neural Networks (RNN) such as disappearing (exploding) gradient, which indicates MSE loss becoming extremely small (large) that significantly distorts the training process of the neural network. As illustrated in Figure 4, the LSTM model addresses this issue by introducing a four-gated mechanism that consists of a forget gate, input gate, update gate, and output gate. This gated mechanism of the model retains long-term FTS data by allowing for circulation within each cell that endorses information circulation within the same LSTM layer and updating the 2<sup>nd</sup> LSTM layer (Figure 4).



**Figure 4.** The structure of an LSTM cell.

Figure 4 illustrates the structure of an LSTM cell (unit).  $C_{t-1}$  stands for the cell's state from the previous temporal step at time  $t-1$ , whereas  $C_t$  stands for the cell's current state at time  $t$ .  $h_{t-1}$  stands for the previous output of the cell within the same LSTM layer that is used as one of the inputs of this illustrated cell at time  $t-1$ ; (upward-going)  $h_t$  stands for the output of this illustrated cell sent to consecutive cell in the same LSTM layer (the 2<sup>nd</sup> LSTM layer) at time  $t$ ;  $x_t$  stands for the input of the cell at the current temporal step at time  $t$ .  $f_t$ ,  $g_t$ ,  $C_t$ , and  $o_t$  indicate the formulas of the forget, input, update, and output gates, respectively.  $\sigma$  indicates the sigmoid function that has a value range between 0 and 1 in which 0 (1) indicates the deletion (retaining) of all information. The  $\tanh$  function, or hyperbolic tangent activation function, normalizes the output data of  $h_t$  by keeping the values within the range of -1 and 1. Capital  $x$  ( $X$ ) (in the forget, update, and output gate) stands for the multiplication operation, and the plus sign ( $+$ ) (in the input gate) stands for the addition operation.

To elaborate the structure of each 4 gates of the model as shown in Figure 4, the forget gate determines the degree of information to be removed for the cell's current state at time  $t$  through the following formula:

$$f_t = \sigma (b_f + W_f x_t + H_f h_{t-1}) \quad (3)$$

where  $\sigma$  is the sigmoid function converting the output value of the function to any value between 0 and 1 in which 0 (1) indicates the deletion (retaining) of all information;  $b_f$  stands for the bias term of the forget gate;  $W_f$  and  $H_f$  stand for the weights of the current input and previous cell output values of  $x_t$  and  $h_{t-1}$ , respectively. Followingly, the model continues with the input gate in which the cell determines the degree of information to be updated for the cell's current state given  $x_t$  and  $h_{t-1}$ . Thus, the formula of the input gate is identical to the forget gate as follows:

$$g_t = \sigma (b_g + W_g x_t + H_g h_{t-1}) \quad (4)$$

Subsequently, the update gate incorporates the first three gates to determine the current cell state value of  $C_t$ . First, the forget gate function is multiplied by the cell's state from the previous temporal step at time  $t-1$ , and the input gate function is multiplied by the update gate function, which is given as follows:

$$\tanh (b_c + W_c x_t + H_c h_{t-1}) \quad (5)$$

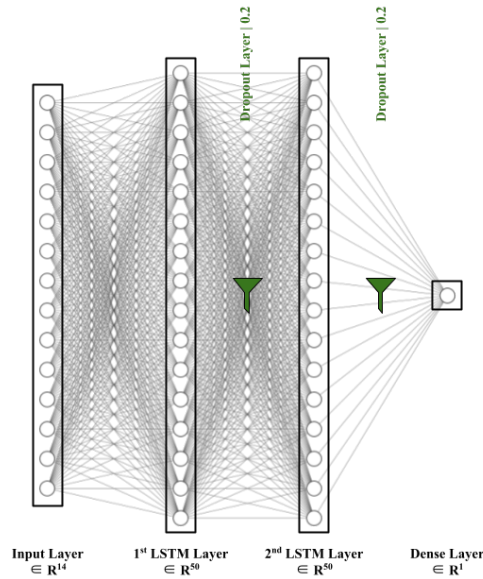
where  $\tanh$  normalizes input values of  $b_c$  and  $x_t$  and  $h_{t-1}$  by keeping the values within the range of -1 and 1. By adding the two multiplication results, the current cell state value is calculated as follows:

$$C_t = f_t * C_{t-1} + g_t * \tanh (b_c + W_c x_t + H_c h_{t-1}) \quad (6)$$

After the current cell state value is set, the LSTM cell lastly calculates its output by multiplying the current cell state firstly by the  $\tanh$  function and secondly by the output gate function, which is also calculated by sigmoid functions of the forget and input gates (Equations 3 and 4). Thus, the final output of the LSTM cell is calculated as follows:

$$h_t = o_t * \tanh (C_t) \quad (7)$$

I used TensorFlow, an open-source library that is commonly used for deep learning models, for the creation of the stacked two-layered LSTM model shown in Figure 5. The window size, given in the input layer, consists of 14 cells, representing the trailing 14 trading days of the price index, so the model processes trailing 14 days of price index data per time step, enabling FTS forecasting. The stacked model consists of two LSTM layers with each having 50 cells, which allows the model to retain long-term and temporal patterns given the considerable size of the price index data. Furthermore, I employ two dropout layers for the training period after each LSTM layer to prevent the model from depending on heavily weighted single or multiple cells and eventually to avoid overfitting the data based on the training data. To this aim, I chose a dropout rate of 0.2 which randomly temporarily removes 20% of the cells in the previous LSTM layer for feeding output values to the next layer. Through this regularization approach in the training data, I aim to increase the validation and prediction performances of the model. Lastly, the dense, or output, layer that consists of 1 cell provides the predicted value of the daily price index (Figure 5).



**Figure 5.** The structure of the stacked two-layered LSTM model with dropout layers.

Figure 5 illustrates the structure of the two-layered LSTM model employed in this study. The input layer consists of 14 cells, representing the last 14 trading days, based on the window size. Both 1st and 2nd LSTM layers consist of 50 cells. The dense layer is the output layer with 1 cell as the desired prediction is on a daily basis for the price index. The dropout layer with a rate of 0.2 is a regularization strategy – *only employed during the training period* – to avoid overfitting the model.

To minimize the training loss, I first normalized the training price index data through the MinMaxScaler formula. As the histogram of the normalized price index data showed values close to 0, I employed the Adam optimizer, which is ideal to use when the data has a sparse characteristic. Besides, I used the MSE loss function to track training and validation losses as this function is suitable for dealing with FTS forecasting. I train the model in 10 epochs where each means the training process is done using the entire training price index data. Yet, I also employed early stopping callback which halts the training process in case there is no progress in the validation loss for 5 consecutive epochs. This early stopping step is a crucial trade-off to avoid overfitting the data by feeding training data if it does not improve the model’s learning performance. Moreover, I chose the batch size of 32, which indicates the model updates the weights (given in the equations from 3 to 6) after every 32 samples of the training data, as it offers an ideal balance between the cell update frequency and computational efforts of the model. After the training and validation processes, the model does out-of-sample daily price index predictions which are based on the weights of the model gained during the training process. Following that, I convert the normalized price index data back to its original scale for the predicted values through a `scaler.inverse_transform` function which is simply an inverse formula of MinMaxScaler. Lastly, I append the out-of-sample daily price index predictions of the model to the same dataset as the actual out-of-sample daily price index to crosscheck prediction accuracy.

### 4.3. Portfolio Performance Metrics

To test the given 4 hypotheses regarding the hybrid portfolio outperforming the S&P Europe 350 Index, the momentum factor, and each respective factor (of) and the entire 5-factor Fama-French model, I employed the following 5 risk-adjusted portfolio performance metrics: Sharpe, Information, Treynor ratios, maximum drawdown (MDD) and alpha. In this regard, the Sharpe ratio is calculated as follows:

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \quad (8)$$

where  $R_p$  refers to the raw monthly returns of the hybrid portfolio strategy;  $R_f$  refers to the risk-free rate based on the one-month treasury bill rate of the U.S.;  $\sigma_p$  refers to the monthly standard deviation of the hybrid portfolio strategy's excess return. I used the Sharpe ratio to present the additional return yielded for additional volatility by investing in this hybrid portfolio strategy relative to the one-month U.S. treasury bill. The greater the Sharpe ratio is, the better the risk-adjusted return of the hybrid portfolio strategy is. Likewise, to compare the portfolio performance of the hybrid strategy with the given benchmarks, I used the Information Ratio.

$$\text{Information Ratio} = \frac{R_p - R_b}{\sigma_{pb}} \quad (9)$$

where  $R_b$  refers to the monthly benchmark returns (S&P Europe 350 Index, the momentum factor, and each respective factor (of) and the entire 5-factor Fama-French model);  $\sigma_{pb}$  refers to the monthly standard deviation of the hybrid portfolio strategy's excess return over respective market benchmarks. Moreover, I used the Treynor Ratio to assess the different market sensitivities of the hybrid portfolio strategy and provide what extent this strategy offsets the given systematic risk.

$$\text{Treynor Ratio} = \frac{R_p - R_f}{\beta_p} \quad (10)$$

where  $\beta_p$  refers to the monthly beta of the hybrid portfolio strategy compared to respective benchmarks. Besides, to identify the risk associated with the hybrid portfolio strategy in terms of significant losses over the out-of-sample period, I used MDD (%), which calculates the greatest highest to the lowest point decline of the portfolio's value until a new peak value is reached again.

$$\text{MDD (\%)} = \frac{\text{Trough Value} - \text{Peak Value}}{\text{Peak Value}} * 100 \quad (11)$$

where the trough value refers to the lowest value of the portfolio until the new highest value is reached; the peak value refers to the highest value of the portfolio before a significant loss. The greater MDD value implies higher significant loss risks associated with the hybrid portfolio strategy. Lastly, I check the excess return of the hybrid portfolio strategy over different respective benchmarks through alpha, which measures the excess returns of a portfolio that cannot be captured by market benchmarks.

$$\alpha = R_p - (R_f + \beta_p * (R_m - R_f)) \quad (12)$$

where  $B_p$  refers to the monthly beta of the hybrid portfolio strategy compared to respective market benchmarks;  $R_m$  refers to the raw monthly returns of respective market benchmarks, so  $R_m$  differentiates across 4 hypotheses based on the benchmark being tested as follows:

Hypothesis 1:  $R_m$  is the monthly return of the S&P Europe 350 Index.

Hypothesis 2:  $R_m$  is the monthly return of the momentum factor.

Hypothesis 3:  $R_m$  is the monthly return of each respective factor of the 5-factor Fama-French model.

Hypothesis 4:  $R_m$  is the monthly return of the whole 5-factor Fama-French model.

Thus, after constructing monthly rebalanced EW long-short hybrid portfolios, I will use the portfolio performance metrics formulas of the Sharpe, Information, Treynor ratios, MDD, and alpha as seen in Equations (8-12) to test the hypotheses. For each hypothesis, I will directly compare the Sharpe ratio and MDD of the hybrid portfolio with those of the given benchmark in a hypothesis. As for the Information ratio, Treynor ratio, and Alpha metrics, I will use the given comparison portfolios in each hypothesis as benchmarks against the hybrid portfolio strategy. In this regard, the benchmark of Hypothesis 1 is the monthly returns of the S&P Europe 350 Index; the benchmark of Hypothesis 2 is the monthly returns of the momentum factor based on the European equity market; the benchmark of Hypothesis 3 is the monthly returns of each respective factor of the 5-factor Fama-French model; the benchmark of Hypothesis 4 is the monthly returns of the whole 5-factor Fama-French model.

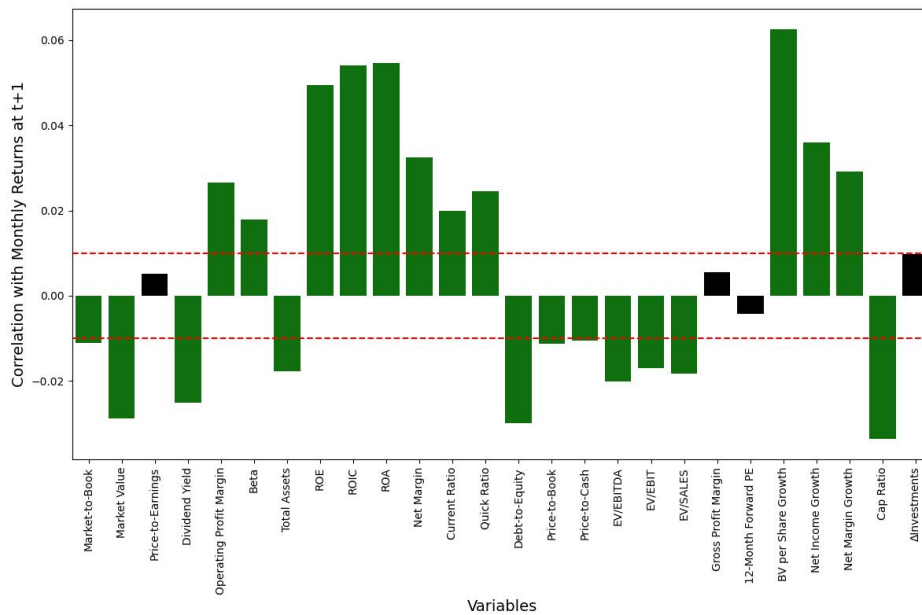
## CHAPTER 5 Results & Discussion

Based on the methodological framework of the hybrid portfolio strategy (Figure 3), this paper first presents the results of the fundamental data section for the creation of the monthly balance sheet score index using the in-sample period in the following order: the validity test, stability test, correlation matrix check (Figure 1A), and PCA analysis (optionally). Followingly, the paper presents the performance of monthly rebalanced quintile EW portfolios sorted on the monthly balance sheet score index on a security level. Subsequently, the paper presents the results of the technical data using the in-sample period in the following order: the creation of the daily price index through PCA analysis and training and validation processes of the LSTM model using the daily price index. Later, the paper presents the results of the out-of-sample price index predictions on a security level, which are used as trading signals. Lastly, the paper presents the performance of the monthly rebalanced long-short EW hybrid portfolio strategy, incorporating monthly 1st (worst-performing) and 5th (best-forming) quintile EW portfolios and trading signals derived from the LSTM model as illustrated in Figure 3.

### 5.1. Fundamental Data

#### 5.1.1 Validity Test

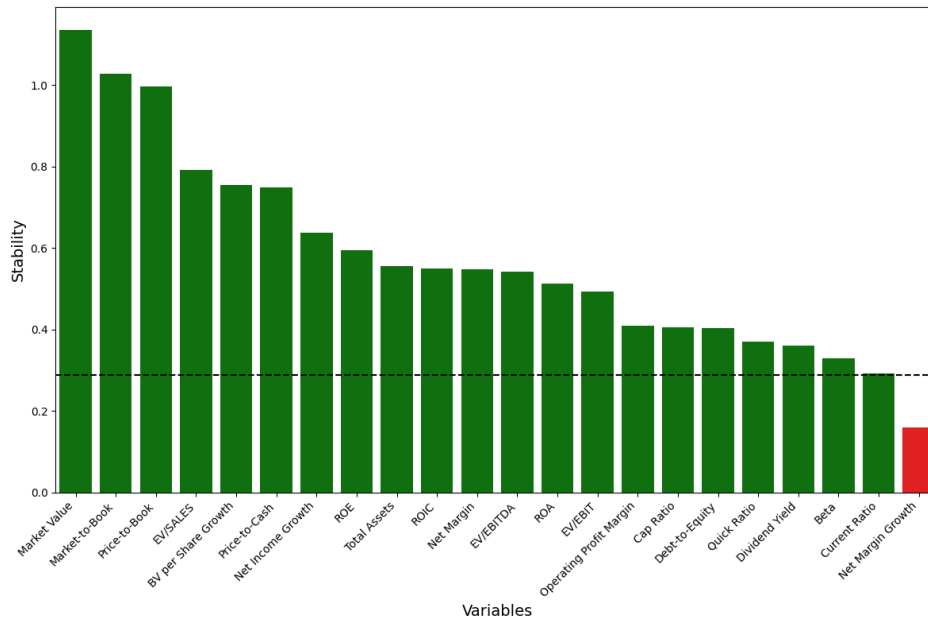
The validity test inspects the information coefficient (IC) of each monthly-basis 26 fundamental data variable during the in-sample period, which identifies the correlation of each variable with the next month's return on a security level. The test aims to eliminate monthly fundamental data variables that do not statistically significantly correlate with the next month's return on a security level as it is essential to identify whether a monthly fundamental data variable has predictive power for the next month's returns on a security level. As displayed by red dotted lines in Figure 6, the threshold of the absolute value of IC values for the fundamental data variables is conservatively set at 0.01, which, as aforementioned, is statistically significant at a 1% level and aimed at preserving the breadth of fundamental data. In this regard, Figure 6 exhibits 11 positively and 11 negatively significantly correlating fundamental data variables with the next month's return on a security level. Among these, book value per share growth, ROE, ROIC, and ROA indicate considerably high positive correlation levels with future returns with IC values above 0.05. Likewise, market value, dividend yield, debt-to-equity, and cap ratio show particularly high negative correlation levels with future returns with IC values below -0.03. Remarkably, Figure 6 points out that monthly growth and profitability-related metrics (Table 1) are significant positive predictors of next month's returns of securities, whilst monthly valuation and solvency-related metrics are significant negative predictors of next month's returns of securities. In Figure 6, the variables highlighted in green are considered valid fundamental variables and are further investigated in the stability test.



**Figure 6.** The IC values of fundamental data variables with the next month’s return on a security level. Figure 6 illustrates the correlation of each monthly fundamental data variable (at time  $t$ ) with the next month’s return (at time  $t + 1$ ) on a security level. The threshold for the IC values is dotted red lines set at both 0.01 and -0.01 since the absolute values of IC values are taken. The black bars indicate no significant correlation level of a fundamental data variable, which has an absolute value of correlation level less than 0.01, whereas the green bars indicate significant correlation levels.

### 5.1.2 Stability Test

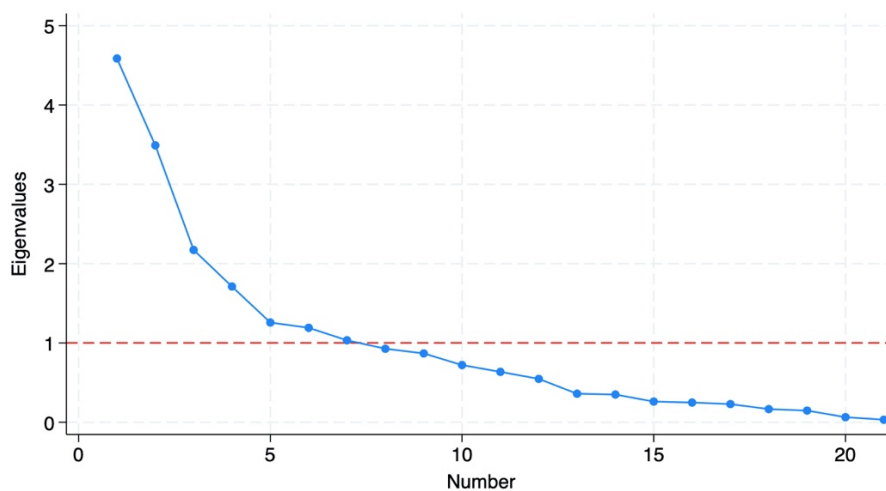
The stability test inspects the information ratio (IR) of 22 valid fundamental variables, that have absolute values of IC higher than, or equal to, 0.01 during the in-sample period (Figure 6), for their respective stability in terms of IC value over the in-sample period to ensure that these valid variables consistently have a statistically significant correlation for the next month’s returns over the in-sample period to be an ideal balance sheet score index contributor over the out-of-sample period. As displayed by black dotted lines in Figure 7, the threshold of the absolute value of IR values for these monthly valid variables is set at 0.289, which, as aforementioned, corresponds to an annualized IR of 1. In this regard, Figure 7 shows that 21 out of 22 valid variables consistently maintain their respective significant IC value, implying stability, over the in-sample period. Among these, only net margin growth (highlighted in red) delivers fairly low stability with an IR value of 0.16, indicating that it is not a reliable predictor for the next month’s returns on a security level. Thus, the valid variables highlighted in green are also considered stable variables, and these core fundamental variables are further investigated in the correlation matrix (Figure 1A) to avoid multicollinearity and if so, to reduce dimensions through PCA analysis.



**Figure 7.** The IR value of valid fundamental data variables (based on their respective IC values). Figure 7 illustrates the stability (IR) of each valid monthly fundamental data variable by dividing the IC value of each valid variable by its respective standard deviation of IC value over the in-sample period. The threshold for the absolute values of IR values is dotted black lines set at 0.289. The green bars indicate stable valid fundamental data variables, whereas the red bar indicates unstable variables.

### 5.1.3 PCA Analysis

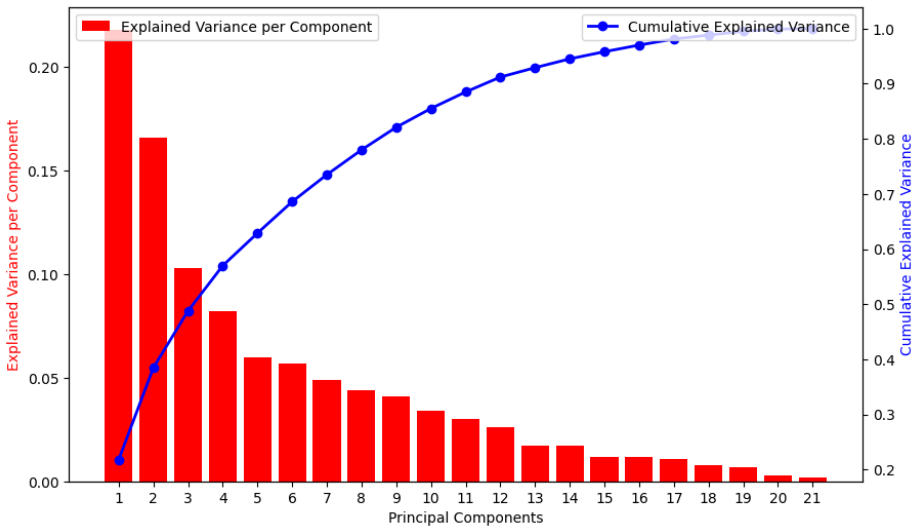
After observing significant correlation levels among 21 core fundamentals in the correlation matrix (Figure 1A), I decided to implement PCA analysis for the creation of a monthly balance sheet score index on a security level to avoid multicollinearity and reduce dimensions, using the Kaiser Rule (i.e. eigenvalue of 1) as mentioned earlier. Based on the scree plot of eigenvalues in Figure 8, principal components up to (including) the 7<sup>th</sup> have eigenvalues greater than 1, and each comprises information *at least* as one respective core fundamental variable. Thus, the balance sheet score index takes only the first 7 principal components into account.



**Figure 8.** The scree plot of eigenvalues of principal components of core fundamentals.

Figure 8 illustrates the eigenvalues of principal components in order from the 1<sup>st</sup> to the 21<sup>st</sup>. The threshold line (y = 1), based on the Kaiser Rule, is the dotted red line. The number (x-axis) shows the principal components in order.

Furthermore, Figure 9 exhibits the cumulative explained variance up to the 7<sup>th</sup> (including) principal component reaching approximately 0.735, implying 73.5% of the total variance of 21 core fundamentals are captured by the first 7 principal components. Along with a concave down increasing cumulative explained variance curve, Figure 9 also shows a downward explained variance per component in which 1<sup>st</sup> component explains approximately 21.8% of the total variance of 21 core fundamentals, whereas 7<sup>th</sup> component explains roughly 4.9% of that variance. Table 1B further informs on the eigenvalues of each component, the difference of eigenvalues of consecutive components, the proportion of total variance explained per component, and cumulative variance across 21 components. Furthermore, the PCA analysis helped to prevent the multicollinearity issue which posed a risk to the accuracy of the balance sheet score index by reducing 21 core fundamentals into 7 principal components which have a covariance of 0. Besides, the factor loadings, or eigenvectors, of the first 7 principal components are provided in Figure 1B. Focusing on the first 3 principal components which explain up to 48.8% of the total variance of 21 core fundamental variables, the factor loadings (Figure 1B) yield interesting insights in which 1<sup>st</sup> principal component indicates significant positive loadings for profitability-related metrics (Table 1) such as ROA, ROIC, ROE, net margin, and operating profit margin; 2<sup>nd</sup> principal component shows considerable positive loadings for valuation and solvency-related metrics (Table 1) such as EV multiples, debt-to-equity, and cap ratio; 3<sup>rd</sup> principal component exhibits not only considerable positive loadings for liquidity-related metrics (Table 1) such as current ratio and quick ratio but also considerable negative loadings for other metrics, such as dividend yield and cap ratio.



**Figure 9.** The combination chart of explained variance per component and their cumulative explained variance of core fundamentals.

Figure 9 illustrates a combination chart of fundamental data in which the proportion of total variance explained by each principal component is shown as red bars, and the cumulative explained variance of principal components is

shown as a blue line. The x-axis line indicates each principal component. The left-side y-axis shows the explained variance per component, and the right-side y-axis shows the cumulative explained variance.

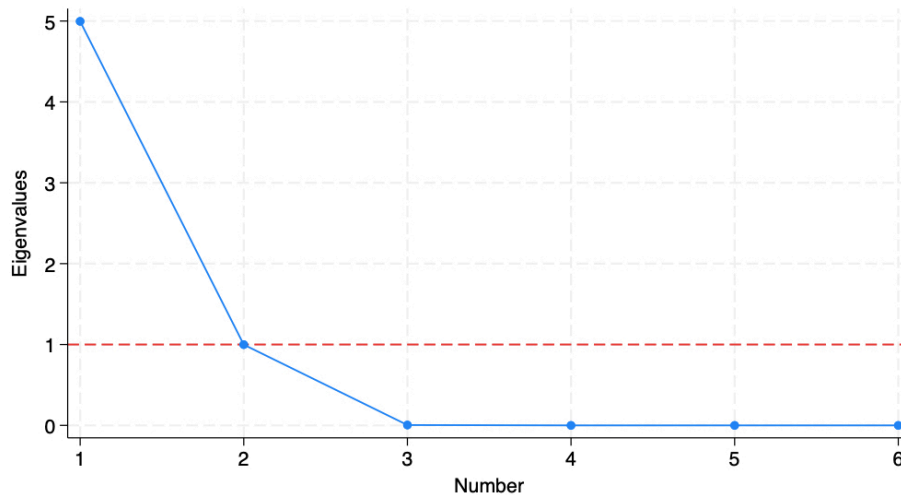
Based on the proportions of total variance explained by each of the 7 principal components for 21 core fundamental variables (Table 1B), I create a monthly balance sheet score index on a security level where these proportions of principal components represent their respective weights for the calculation of the balance sheet score index. Figure 2B presents the histogram of this index, which has a positive mean yet close to zero and slight skewness towards the left. Furthermore, the monthly balance sheet score index is calculated as follows:

$$\begin{aligned} \text{Balance Sheet Score Index} = & 0.218 * PC1 + 0.166 * PC2 + 0.103 * PC3 \\ & + 0.082 * PC4 + 0.060 * PC5 + 0.057 * PC6 + 0.049 * PC7 \end{aligned} \quad (13)$$

## **5.2. Technical Data**

### **5.2.1 PCA Analysis**

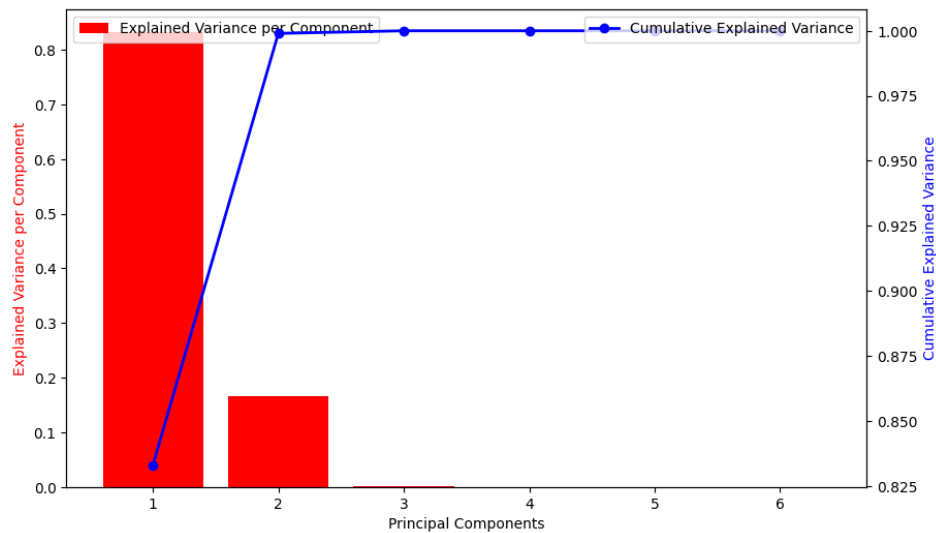
Similar to the fundamental data, OHLCAV variables also display high pairwise correlation levels, excluding daily volume levels of securities, as shown in Table 2B. These pairwise correlations are perfect (equal to 1) among daily opening, high, low, and closing prices of securities, and they are each almost perfectly correlated (equal to 0.982) with daily adjusted closing prices. Conversely, the daily volume of securities indicates much lower and negative correlation levels with other price variables, ranging between -0.067 and -0.078. The high pairwise correlation levels among 6 daily trading volume and price variables imply the need for PCA analysis which would not only reduce dimensions but also lower the computational efforts when running the LSTM model. Implementing PCA analysis and deriving a daily price index (on a security level) would allow feeding the LSTM model with 1694840 daily trading volume and price information instead of 10169040 daily information overall and thus vastly lower the redundant computational efforts of the model. Furthermore, the pairwise correlation given in Table 2B hints these variables could be potentially reduced to 2 principal components, which later could be used to create the daily price index on a security level with the respective weights of each principal component. Thus, to avoid multicollinearity and reduce dimensions, I implemented PCA analysis, using the Kaiser Rule (i.e. eigenvalue of 1). Based on the scree plot of eigenvalues in Figure 10, principal components up to (including) the 2nd have eigenvalues greater than, or equal to, 1, and each comprises information at least as one respective OHLCAV variable. Thus, the daily price index takes only the first 2 principal components into account, which supports the findings of the pairwise correlations shown in Table 2B.



**Figure 10.** The scree plot of eigenvalues of principal components of OHLCAV variables.

Figure 10 illustrates the eigenvalues of principal components in order from the 1<sup>st</sup> to the 6<sup>th</sup>. The threshold line ( $y = 1$ ), based on the Kaiser Rule, is the dotted red line. The number (x-axis) shows the principal components in order.

Furthermore, Figure 11 illustrates the cumulative explained variance up to the 2<sup>nd</sup> (including) principal component reaching approximately 0.999, suggesting 99.9% of the total variance of 6 OHLCAV variables are captured by the first 2 principal components. Along with a steep cumulative explained variance curve in the first two components, Figure 11 also shows a sharp declining explained variance per component in which 1<sup>st</sup> component explains approximately 83.3% of the total variance of OHLCAV variables, whereas 2<sup>nd</sup> component explains roughly 16.6% of that variance. Table 3B further informs on the eigenvalues of each component, the difference of eigenvalues of consecutive components, the proportion of total variance explained per component, and cumulative variance across 6 components. Moreover, the PCA analysis helped to prevent the multicollinearity issue of the OHLCAV dataset which posed a risk to the accuracy of the balance sheet score index by reducing 6 variables into 2 principal components which have a covariance of 0. Besides, the factor loadings, or eigenvectors, of the first 2 principal components are provided in Figure 3B, which shows the ideal dimension reduction. The 1<sup>st</sup> principal component indicates significant positive loadings for daily opening, highest, lowest, closing, and adjusted closing prices variables (each having an eigenvector of 0.447); therefore, the 1<sup>st</sup> principal component could be described as the daily price component. Besides, the 2<sup>nd</sup> principal component indicates significant positive loading for the daily volume variable (having an eigenvector of 1); therefore, the 2<sup>nd</sup> principal component could be described as the daily volume component.



**Figure 11.** The combination chart of explained variance per component and their cumulative explained variance of OHLCAV variables.

Figure 11 illustrates a combination chart of technical data in which the proportion of total variance explained by each principal component is shown as red bars, and the cumulative explained variance of principal components is shown as a blue line. The x-axis line indicates each principal component. The left-side y-axis shows the explained variance per component, and the right-side y-axis shows the cumulative explained variance.

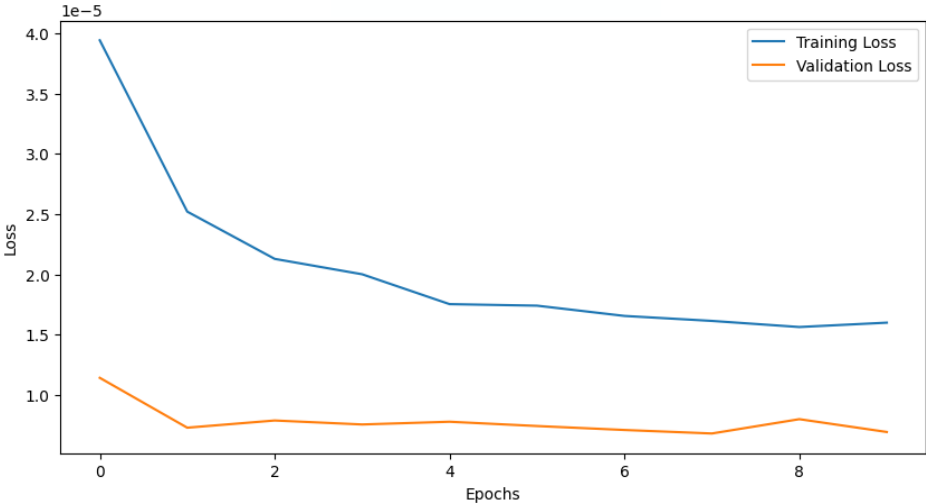
Based on the given proportions of total variance explained by each of the first 2 principal components for 6 daily trading volume and price variables (Table 3B), I create a monthly daily price index on a security level where these proportions of principal components represent their respective weights for the calculation of the price index. Thus, the daily price index is calculated as follows:

$$Price\ Index = 0.833 * PC1 + 0.166 * PC2 \quad (14)$$

### 5.2.2 LSTM Neural Network

After calculating the daily price index on a security level over the in-sample period, which is based on the given proportions of 2 principal components as seen in Equation (14), the normalization of price index data is implemented through the MinMaxScaler formula due to the sparse characteristic of the price index over a wide array of securities. The normalization helps the LSTM model to effectively minimize the training loss. Illustrating the training and validation mean squared error (MSE) loss of the model over 10 epochs, Figure 12 exhibits the validation loss to be consistently lower than the training loss despite their convergence over epochs. The use of dropout layers only being applied during the training process can explain this discrepancy. The dropout layer with a rate of 0.2 as a regularization strategy is applied to avoid overfitting the model based on the learnings of the training data. Thus, the dropout layer further complicates the training process by yielding additional regularization loss on the training process. Furthermore, Figure 12 exhibits a steep decline in training loss in the early phases of training, particularly between the 1st and 4th epochs, and a relatively stable loss with a slight decline

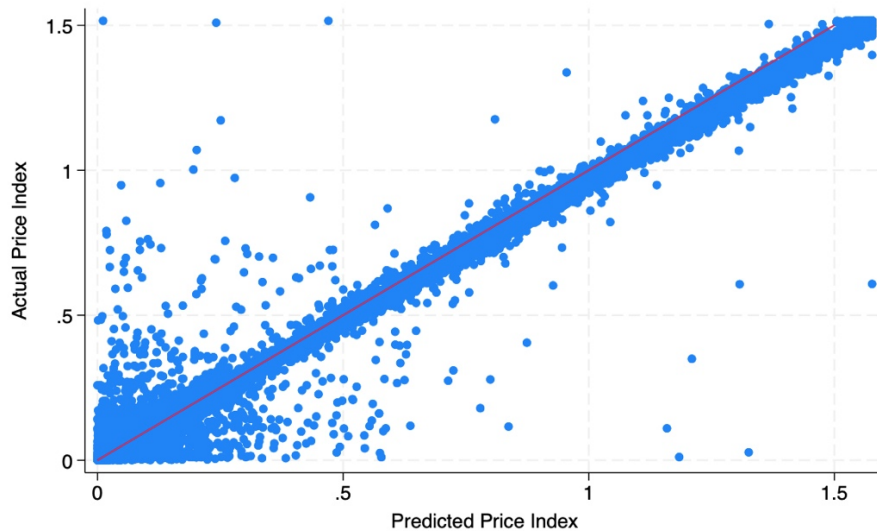
until the 10th epoch. As for the validation data, the model shows a relatively stable loss curve over 10 epochs after a slight decline in validation loss in 1st epoch (Figure 12). Nevertheless, this stable loss curve of validation data did not activate the early stopping callback which would halt the training process in case there was no progress in the validation loss for 5 consecutive epochs. Lastly, Figure 12 indicates a good balance between the learning of the model through declining training loss and the generalization of the model through a relatively low and stable validation loss.



**Figure 12.** The training and validation MSE loss of the LSTM model over 10 epochs.

Figure 12 illustrates the change of the training and validation mean squared error (MSE) loss over 10 epochs. Only during the training process, the dropout layer with a rate of 0.2 as a regularization strategy is applied to avoid overfitting the model, while the validation process has no dropout layer. Thus, the higher levels of training loss could be explained by the conditional dropout layers which further complicate the training process by randomly dropping 20% of cells at each epoch.

Following this, the model provided out-of-sample daily price index predictions on a security level which are based on the weights of the model gained during the training process. To identify the degree of alignment of the model predictions concerning the actual data over the out-of-sample period, I created an actual price index which is also calculated as shown in Equation (14) based on the actual daily OHLCAV data of securities. As displayed in the scatter plot of the actual and predicted price index (Figure 13), there is an apparent positive linear relationship between both as the majority of observations are distributed around the 45-degree line. This positive linear relationship is further pronounced in the *almost perfect* correlation between these two indexes accounting for 0.999 as shown in Table 4B. Despite the overall satisfactory prediction accuracy of the LSTM model, there are extreme outliers towards both ends of the index level, implying a potential overfitting, or underfitting, of the model.



**Figure 13.** The scatter plot of actual and predicted price index.

Figure 13 illustrates the scatter plot of the actual price index, which is calculated using Equation (14) based on the actual daily data of OHLCAV variables, and the predicted price index, which is derived from daily forecasts of the LSTM model, over the out-of-sample period. The red line provides a 45-degree line to illustrate a linear positive relationship between the actual and predicted price index. Due to their sparse characteristics, the given values of actual and predicted price indexes are standardized for illustration purposes, and later, the predicted price index is converted back to its original scale for portfolio construction.

### **5.3. Portfolio Performance**

The first stage of the hybrid portfolio strategy was to create monthly rebalanced EW quintiles that were sorted on the monthly balance sheet score index (derived from core fundamentals). Thus, it is crucial to analyze the performance of these quintiles to identify whether the first stage was accomplished by identifying and differentiating the worst and best-forming securities in terms of returns based on fundamental data. Table 3 provides the relevant performance metrics of quintiles, where the first quintile Q1 (fifth quintile Q5) represents the worst-performing (best-forming) securities based on the monthly balance sheet score index over the out-of-sample period. In this regard, the monthly average rate of returns across quintiles is sharply and consistently increasing from Q1 to Q5 in which Q1 and Q5 yield 6 and 108 basis points (bps) per month, respectively. While the returns are higher, on average, towards higher quintiles, the monthly standard deviations seem to be stable across quintiles with Q1 having the highest monthly standard deviation of 0.06. Maximum drawdown consistently declines from Q1 (-33.42%) to Q4 (-20.42%); however, there is a significant increase observed in Q5 (-26.73%), implying the greater on average monthly returns are accompanied by potentially greater significant losses relative to Q2-Q4. Similar to ARR, the Sharpe ratio is consistently improving from Q1 to Q5, suggesting Q5 with higher on-average returns is also efficient on a risk-adjusted level and generates 0.2 units of excess return for each unit of additional risk taken, relative to the risk-free rate. Lastly, the skewness levels across quintiles range between -0.33 and 0.10, suggesting that the monthly returns distributions of

quintiles are fairly symmetrical and monthly returns of higher quintiles are relatively left skewed. Overall, Table 3 presents that a higher balance sheet score index not only yields higher monthly returns on average but also offers better risk-adjusted performance.

**Table 3.** Performance metrics of monthly fundamental data-based EW quintiles.

| Metrics  | Q1      | Q2      | Q3      | Q4      | Q5      |
|----------|---------|---------|---------|---------|---------|
| ARR      | 6       | 64      | 69      | 84      | 108     |
| SD       | 0.06    | 0.05    | 0.05    | 0.05    | 0.05    |
| MDD      | -33.42% | -24.95% | -22.51% | -20.42% | -26.73% |
| Sharpe   | -0.01   | 0.10    | 0.12    | 0.15    | 0.20    |
| Skewness | 0.10    | -0.26   | -0.31   | -0.28   | -0.33   |

Table 3 provides the performance metrics of quintiles, where the first quantile Q1 (fifth quantile Q5) represents the worst-performing (best-forming) securities, for the out-of-sample period. These monthly rebalanced EW quintiles are sorted on the monthly balance sheet score index. ARR stands for the monthly average rate of return provided in basis points (bps); SD stands for the monthly standard deviation; MDD (maximum drawdown) and Sharpe ratio are calculated as seen in Equation (11) and Equation (8), respectively; Skewness checks for the asymmetrical distribution of monthly returns in each respective quintile.

After incorporating trading signals of the LSTM network based on the daily price index, I constructed the final monthly rebalanced EW long-short hybrid portfolio in which I take a long position in the top 10% of Q5 securities, while I take a short position in the bottom 10% of Q1 securities both based on the forecasted monthly return of daily price index. In this regard, Table 4 provides the comparative portfolio performance of this hybrid strategy across different benchmarks as mentioned in Hypotheses 1-4, in which the Information ratio, Treynor ratio, and alpha metrics of the hybrid portfolio strategy are provided under the corresponding benchmark. The hybrid portfolio has a notably greater annualized Sharpe ratio of 1.45 relative to other benchmarks, suggesting efficient risk-adjusted performance of the portfolio. To compare whether this superior risk-adjusted performance of the hybrid portfolio strategy persists after directly comparing its performance against given benchmarks through the annualized Information ratio (Table 4), I found the hybrid strategy moderately to significantly outperforms all benchmarks (with IR value ranging between 0.87 and 1.61) compared to standard deviation of excess returns of the hybrid portfolio over these benchmarks. The Sharpe and Information ratios in Table 4 offer the risk-adjusted outperformance of the hybrid portfolio against the risk-free rate, S&P Europe 350 Index, and other factors. However, the annualized Treynor ratio has relatively more mixed but heavily negative results in which the hybrid portfolio, for instance, has a significant negative value of -4.08 when benchmarked against the market risk premium factor (Mkt-RF). Given that the average rate of returns (ARR) of the hybrid portfolio is 294 bps per month (Table 4), these negative Treynor ratios suggest that the hybrid portfolio has a negative beta with market factors instead of

performing worse than the risk-free rate, which would not align with the given results of Sharpe and Information ratios. This negative beta can be explained to some extent by the short leg that accounts for 50% of the hybrid portfolio. Furthermore, these varying annualized Treynor ratio values in Table 4 suggest that the hybrid portfolio is either negatively related or unrelated to systematic risks; therefore the hybrid portfolio is not offsetting for the positive market risks.

**Table 4.** Performance metrics of benchmarks and hybrid portfolio strategy.

| Metrics      | Index   | Mkt-RF  | SMB     | HML     | RMW     | CMA     | WML     | 5-FF    | Portfolio |
|--------------|---------|---------|---------|---------|---------|---------|---------|---------|-----------|
| Sharpe       | 0.33    | 0.32    | -0.42   | 0.04    | 0.16    | -0.52   | 0.25    | 0.06    | 1.45      |
| IR*          | 0.95    | 0.87    | 1.61    | 0.98    | 1.60    | 1.33    | 1.43    | 1.31    | -         |
| Treynor*     | -0.59   | -4.08   | 0.12    | -0.08   | 0.04    | -0.04   | 0.13    | -0.07   | -         |
| MDD          | -23.03% | -31.54% | -21.70% | -29.56% | -6.69%  | -27.20% | -19.87% | -10.02% | -16.74%   |
| Alpha*       | 0.40*** | 0.39*** | 0.42*** | 0.40*** | 0.37*** | 0.29*** | 0.36*** | 0.41*** | -         |
| ARR<br>(bps) | 59      | 68      | -9      | 19      | 22      | -16     | 42      | 84      | 294       |
| SD           | 0.05    | 0.06    | 0.02    | 0.04    | 0.02    | 0.02    | 0.04    | 0.08    | 0.07      |
| Skew         | -0.33   | -0.14   | 0.17    | 0.35    | -0.59   | 0.04    | -2.18   | -0.98   | 0.55      |

Table 4 provides the portfolio performance metrics across different benchmarks and the hybrid portfolio strategy. The index represents for the S&P Europe 350 Index; Mkt-RF represents the market risk premium factor; SMB represents the size factor; HML represents the value factor; RMW represents the profitability factor; CMA represents the investment factor; WML represents the momentum factor; 5-FF represents the five-factor Fama-French model, which is a combination of Mkt-RF, SMB, HML, RMW, and CMA factors; Portfolio represents the hybrid portfolio strategy. All returns are on a monthly basis. All factor returns are retrieved from the Kenneth R. French data library. The index returns are retrieved from the FactSet database. \*\* and \*\*\* indicate significance at %5 and %1 levels, respectively. For the Information ratio, Treynor Ratio, and Alpha metrics, the given portfolios are used as benchmarks against the hybrid portfolios strategy. Sharpe, Information, and Treynor ratios are annualized, while alpha is annualized with compounding. Sharpe, Information, Treynor, MDD (maximum drawdown), and Alpha metrics are calculated as seen in Equations (8-12). ARR (monthly average rate of returns) is provided in basis points (bps); SD stands for the monthly standard deviation; Skewness checks for the asymmetrical distribution of monthly returns in each respective portfolio.

*\*Note.* Information ratio (IR), Treynor ratio, and Alpha metrics are not provided separately for Portfolio (hybrid portfolio strategy) as these metrics are calculated relative to each benchmark and presented under the corresponding benchmark.

Followingly, Table 4 presents the hybrid portfolio's maximum drawdown (MDD) of -16.74% across other benchmarks over the out-of-sample period is fairly lower relative to other benchmarks such as the S&P Europe 350 Index, and the profitability factor and 5-Factor Fama-French model have notably low MDD values of -6.69% and -10.02%, respectively. Figure 1D-9D further illustrates MDD values with their respective peak and trough points. Additionally, Table 4 provides that the annualized alpha

(with compounding) is statistically significant across each given benchmark at a 1% level. This suggests that the monthly returns of the hybrid portfolio cannot be explained by the monthly returns of given benchmarks, and the hybrid portfolio, as a result, generates statistically significant positive abnormal returns over these benchmarks. Table 1C provides detailed time-series regressions of hybrid portfolio strategy across the given market benchmarks on a monthly basis regarding the alpha. Furthermore, the portfolio descriptive statistics in Table 4 show that the significant monthly average rate of return of 294 bps of the hybrid portfolio is remarkably greater than other market benchmarks and is not accompanied by excessive standard deviation (0.07) relative to the other market benchmarks. Lastly, the hybrid portfolio's positive skewness of 0.55 with its given positive ARR of 294 bps offers there are substantial positive monthly returns for the portfolio. Conversely, the significantly negative skewness of the momentum factor (WML) implies a potential momentum reversal over the out-of-sample period. Overall, the hybrid portfolio strategy is found to be significantly outperforming other market benchmarks given Sharpe, Information, and Alpha metrics; however, negative Treynor ratios are indicative of the hybrid portfolio's shortcoming of not offsetting (positive) market risks, particularly for the short-leg of the portfolio.

Therefore, the results displayed in Table 4 indicate that hypothesis 1 is accepted based on the Sharpe ratio, Information ratio, MDD, and alpha metrics as the hybrid portfolio strategy significantly outperforms the S&P Europe 350 Index based on these metrics; however, negative Treynor ratio of the hybrid portfolio for the index implies that the hybrid portfolio strategy may not offset the positive market risks associated with the index. Likewise, hypothesis 2 is accepted based on all given metrics of Sharpe ratio, Information ratio, Treynor ratio, MDD, and alpha metrics as the hybrid portfolio strategy significantly outperforms the momentum factor on these metrics. Hypothesis 3 is accepted based on the Sharpe ratio, Information ratio, and alpha metrics; however, the lower MDD value of the profitability factor and negative Treynor ratios of the hybrid portfolio for the market risk premium, value, and investment factors imply that the hybrid portfolio strategy may not offset the positive market risks and is relatively riskier than the profitability factor based on MDD. Lastly, hypothesis 4 is accepted based on the Sharpe ratio, Information ratio, and alpha metrics as the hybrid portfolio strategy significantly outperforms the 5-factor Fama-French model based on these metrics; however, the lower MDD value of the 5-factor model and negative Treynor ratio of the hybrid portfolio for the 5-factor model implies that the hybrid portfolio strategy may not offset the positive market risks and is relatively riskier than the 5-factor model based on MDD.



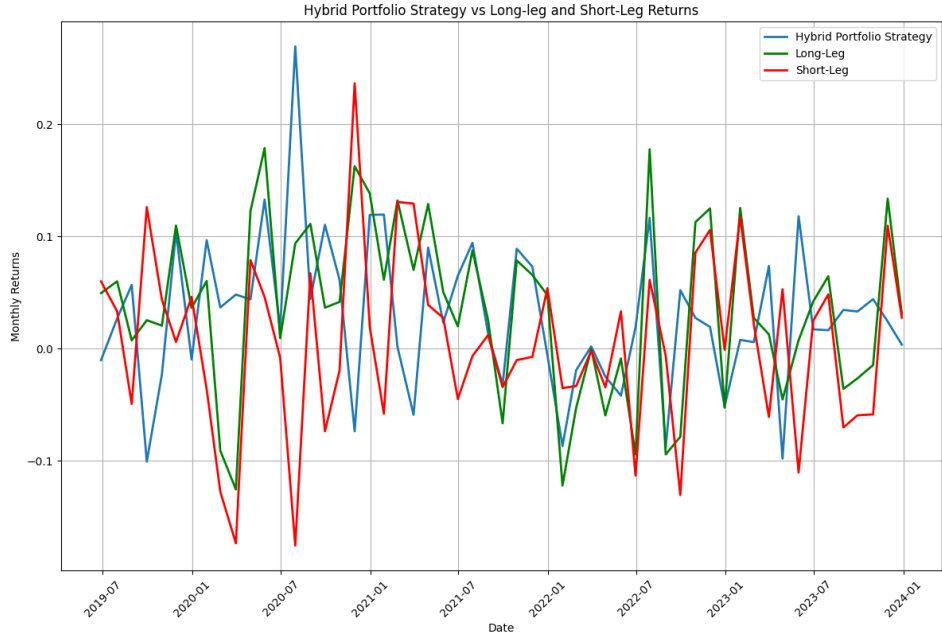
**Figure 14.** The comparative portfolio performance of hybrid portfolio strategy across hypotheses.

Figure 14 illustrates the comparison of monthly returns of benchmarks across the given 4 hypotheses with those of the hybrid portfolio strategy in 4 sub-figures. Sub-figure Hypothesis 1 compares the S&P Europe 350 Index against the hybrid portfolio strategy; Sub-figure Hypothesis 2 compares WML (momentum factor) against the hybrid portfolio strategy; Sub-figure Hypothesis 3 compares Mkt-RF (market size premium), SMB (size), HML (value), RMW (profitability), and CMA (investment) factors against the hybrid portfolio strategy; Sub-figure Hypothesis 4 compares the 5-factor Fama-French model against the hybrid portfolio strategy.

Given the insights of Table 4, Figure 14 illustrates the performance of the hybrid portfolio strategy in terms of raw monthly returns in 4 sub-figures across different market benchmarks in Hypotheses 1-4. In the Hypothesis 1 sub-figure of Figure 14, the hybrid portfolio has a varying performance relative to the S&P Europe 350 Index; however, local peaks and troughs of the Hypothesis 1 sub-figure suggest that the hybrid portfolio strategy can often identify the bidirectional price movements of the index sufficiently. Followingly, the Hypothesis 2 sub-figure of Figure 14 shows similar directional movements between the hybrid portfolio strategy and the momentum factor (WML), while the magnitude of these similar directional movements is on average greater in the hybrid portfolio strategy, implying the LSTM model possibly being efficient at selecting and weighting momentum signals. Moreover, the Hypothesis 3 sub-figure of Figure 14 exhibits somewhat more mixed results between each factor of the 5-factor model and the hybrid portfolio strategy. Notably, there is a potential lagged correlation between the hybrid portfolio strategy and the market risk premium (Mkt-RF) in which the sharp decline or increase in monthly returns of the hybrid portfolio is later reflected in Mkt-RF.

Lastly, the Hypothesis 4 sub-figure of Figure 14 presents the 5-factor model to be less volatile relative to the hybrid portfolio, while the hybrid portfolio seems to outperform the 5-factor model, on average.

To investigate further from which leg the monthly returns of the long-short hybrid portfolio originate, Figure 15 illustrates the comparison between the long and short-leg returns of the portfolio. In Figure 15, the lower returns in the short leg and greater returns in the long leg imply the better performance of the hybrid portfolio strategy. Despite the high volatility of both legs, there is no clear outperformance of any. Furthermore, Figure 15 presents that in the months that the short leg or long leg underperformed, it was compensated by the other leg to prevent substantial losses, suggesting the efficient risk management of the portfolio.



**Figure 15.** The long and short-leg returns of the hybrid portfolio strategy.

Figure 15 illustrates the comparison of monthly returns of long-leg (highlighted in green) and short-leg (highlighted in red) of the hybrid portfolio strategy.

## CHAPTER 6 Conclusion

In this thesis, I aimed to answer the following research question: *how does the hybrid portfolio strategy, using fundamental sorting and LSTM Networks for technical indicators, perform relative to the S&P Europe 350 Index, momentum factor, and 5-factor Fama-French model?* The objective of this study is to pursue a consistent abnormal return-generating investment strategy on a risk-adjusted basis through devising a hybrid portfolio strategy that integrates principal component analysis (PCA) both for fundamental factor sorting and for Long Short-Term Memory (LSTM) neural networks for the technical data and using the investible universe of the S&P Europe 350 Index constituents. In this regard, as illustrated in a diagram in Figure 3, I first inspected the stability and validity of fundamental data to create a monthly balance sheet score index, for which I found 21 out of 26 balance sheet material-related variables to be both valid and stable over the in-sample period. Thus, I decided to employ PCA analysis for fundamental data in addition to the technical data to reduce dimensions, avoid multicollinearity, and lower the computational efforts for the LSTM model. After deriving the daily price index and monthly balance sheet score index on a security level by PCA analysis, I used the LSTM model to train and validate the daily price index to do out-of-sample prediction which was later used as trading signals. Similarly, I sorted securities based on the monthly balance sheet score index to construct monthly rebalanced EW quintiles. Lastly, I constructed the monthly rebalanced EW long-short hybrid portfolio by integrating the balance sheet score-based quintiles and trading signals of the LSTM network by taking a long position in the top 10% of Q5 (best-performing) securities, while I took a short position in the bottom 10% of Q1 (worst-performing) securities based on trading signals.

To investigate the performance aligned with Hypotheses 1-4, I analyzed the performance of the hybrid portfolio strategy relative to the S&P Europe 350 Index, the momentum factor, and each factor of the 5-factor Fama-French model as well as the complete 5-factor model. As a result, I found that this hybrid strategy significantly outperforms all given market benchmarks for annualized Sharpe and annualized Information ratios, in which the Sharpe ratio of the hybrid portfolio and the greatest relevant market benchmark (the S&P Europe 350 Index) were 1.45 and 0.33, respectively. Similarly, the annualized Information ratio of the hybrid portfolio when benchmarked against market factors and the index was found to range between 0.87 (Mkt-RF) and 1.61 (SMB). Followingly, the hybrid portfolio had the third lowest MDD value of 16.74%. Despite this substantial loss, it still had a lower MDD level compared to the other 6 market benchmarks which are diversified more. Conversely, the hybrid portfolio was found to have annualized Treynor ratio values being either negative or close to zero with the emphasis on the Treynor ratio of the portfolio reaching -4.08 when benchmarked against the market risk premium factor (Mkt-RF), suggesting that this strategy does not offset the positive market risk in overall. Furthermore, the annualized alpha (with compounding) of the hybrid portfolio was found to be statistically significant across each given benchmark at varying 1% to 5% levels. This suggests that the monthly returns of the hybrid portfolio cannot be explained by the monthly returns of given benchmarks,

and the hybrid portfolio, as a result, generates statistically significant positive abnormal returns over these benchmarks. Thus, I accepted all 4 hypotheses as results in Table 4 indicate a statistically significant outperformance of the hybrid portfolio strategy over all given market benchmarks in risk-adjusted returns (Table 5).

**Table 5.** Summary of hypothesis test results of the hybrid portfolio strategy.

| # Hypothesis | Sharpe | Information | Treynor | MDD | Alpha |
|--------------|--------|-------------|---------|-----|-------|
| 1            | +      | +           | -       | +   | +     |
| 2            | +      | +           | +       | +   | +     |
| 3            | +      | +           | +/-     | +/- | +     |
| 4            | +      | +           | -       | -   | +     |

Table 5 provides the information on whether each respective hypothesis for each portfolio performance metrics is accepted or rejected as given + and -, respectively. +/- stands for the mixed results as the results for each respective factor in hypothesis 3 differentiates for Treynor ratio and maximum drawdown (MDD) values.

Furthermore, the integration of the PCA analysis and LSTM network was efficient in feeding the model with relevant (valid and stable) fundamental and technical data that was free of multicollinearity and each principal components were weighted in their respective index (balance sheet score or price) calculations based on their respective proportion of variance explanation. The monthly rebalanced EW quintiles sorted on balance sheet score indexes of securities confirm the efficiency of this process by the Q5 quintile significantly outperforming the Q1 quintile. Q5 was found to yield 108 bps per month with a standard deviation of 0.05 and Sharpe ratio of 0.2, whereas Q1 was found to yield 6 bps per month with a standard deviation of 0.06 and Sharpe ratio of -0.01. This performance gap between Q1 and Q5 was further extended by the trading signals of the LSTM network in which the hybrid portfolio accomplished generating 294 bps per month, demonstrating the contribution of both PCA analysis and the LSTM model in this hybrid strategy.

However, the optimistic outlook and findings of this paper regarding the hybrid portfolio strategy need to be approached with vigilance due to several limitations. Firstly, the parameter settings of the LSTM network were chosen based on the academic literature, and advanced optimization techniques were not employed, which could pose a risk for the model not performing this well in other time and investible universe settings. Additionally, even though the hybrid portfolio strategy was free of survivorship and look-ahead biases, the monthly rebalanced EW long-short portfolio consisted of only 12 securities each month in which the long-leg and short-leg both included 6 securities. Each security having 8.33% of the total monthly portfolio value can imply a low diversification level which could be improved by using a relatively wider investible universe. Lastly, this study omitted short-selling constraints and transaction costs associated with the active investing strategy of the hybrid portfolio. Thus, overlooking these constraints, particularly for retail investors, could lead to an overestimation of the (monthly) returns that are presented in this hybrid portfolio strategy.

## REFERENCES

- Alexander, S. S. (1961). Price movements in speculative markets: Trends or random walks. *Industrial Management Review (pre-1986)*, 2(2), 7.
- American Economic Association. (2020). Processing the rise of quantitative investing. *American Economic Association*. <https://www.aeaweb.org/research/charts/financial-technology-big-data-information>
- Asness, C. S., Liew, J. M., & Stevens, R. L. (1997). Parallels between the cross-sectional predictability of stock and country returns. *Journal of Portfolio Management*, 23(3), 79.
- Bachelier, L. (1964). Theory of speculation. In P. Cootner (Ed.), *The random character of stock market prices* (pp. 17-78). Cambridge, MA: MIT Press. (Original work published 1900).
- Bagnara, M. (2024). Asset pricing and machine learning: A critical review. *Journal of Economic Surveys*, 38(1), 27-56.
- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3-18. [https://doi.org/10.1016/0304-405X\(81\)90018-0](https://doi.org/10.1016/0304-405X(81)90018-0)
- Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of Behavioral Economics: Applications and Foundations 1* (pp. 79-175). Elsevier, North-Holland.
- Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The Journal of Finance*, 32(3), 663-682.
- Bender, J., Briand, R., Nielsen, F., & Stefek, D. (2010). Portfolio of risk premia: A new approach to diversification. *Journal of Portfolio Management*, 36(2), 17-25. <https://doi.org/10.3905/jpm.2010.36.2.017>
- Bettman, J. L., Sault, S. J., & Schultz, E. L. (2009). Fundamental and technical analysis: substitutes or complements? *Accounting & Finance*, 49(1), 21-36.
- BlackRock, Inc. (2024). Guide to investing in long/short funds. BlackRock. Retrieved May 15, 2024, from <https://www.blackrock.com/us/individual/products/long-short-funds>
- Botosan, C. A., & Plumlee, M. A. (2005). Assessing alternative proxies for the expected risk premium. *The Accounting Review*, 80(1), 21-53. <https://doi.org/10.2308/accr.2005.80.1.21>
- Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business*, 78(2), 405-440. <https://doi.org/10.1086/427633>
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.
- Chandrashekar, V. (2015). The Evolution of Quantitative Investment Strategies. American Century Investments. Retrieved from <https://www.pionline.com/assets/docs/CO98933327.PDF>
- Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. In *2015 IEEE International Conference on Big Data (Big Data)*

- (pp. 2823-2824). IEEE.
- Clarke, R., de Silva, H., & Thorley, S. (2016). Fundamentals of efficient factor investing (corrected May 2017). *Financial Analysts Journal*, 72(6), 9–26. <https://doi.org/10.2469/faj.v72.n6.3>
- Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, 66(4), 1047-1108.
- de Prado, M. M. L. (2020). *Machine learning for asset managers*. Cambridge University Press.
- Dixon, M., Klabjan, D., & Bang, J. H. (2017). Classification-based financial markets prediction using deep neural networks. *Algorithmic Finance*, 6(3-4), 67-77.
- Fabozzi, F. J., Focardi, S. M., & Kolm, P. N. (2010). *Quantitative equity investing: Techniques and strategies*. John Wiley & Sons.
- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1), 34-105.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383–417. <https://doi.org/10.2307/2325486>
- Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427-465.
- Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1), 55-84.
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.
- Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), 607-636.
- Fang, Z., Ma, X., Pan, H., Yang, G., & Arce, G. R. (2023). Movement forecasting of financial time series based on adaptive LSTM-BN network. *Expert Systems with Applications*, 213, 119207.
- Fang, Zhong, Y., Xie, C., Zhang, H., & Li, S. (2020). Research on PCA-LSTM-based short-term load forecasting method. *IOP Conference Series: Earth and Environmental Science*, 495(1), 12015. <https://doi.org/10.1088/1755-1315/495/1/012015>
- Farboodi, M., & Veldkamp, L. (2020). Long-run growth of financial data technology. *American Economic Review*, 110(8), 2485-2523.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669.
- Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. *Int. J. Intell. Technol. Appl. Stat*, 11(2), 105-111.
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222-2232.
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223-2273.

- Hanauer, M. X. (2020). A comparison of global factor models (SSRN Scholarly Paper 3546295). <https://doi.org/10.2139/ssrn.3546295>
- Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the cross-section of expected returns. *The Review of Financial Studies*, 29(1), 5-68.
- Haugen, R. A., & Heins, A. J. (1975). Risk and the rate of return on financial assets: Some old wine in new bottles. *Journal of Financial and Quantitative Analysis*, 10(5), 775-784.
- Heaton, J. B., Polson, N. G., & Witte, J. H. (2016). Deep learning in finance. *arXiv preprint arXiv:1602.06561*.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies*, 33(5), 2019-2133.
- Huij, J., Lansdorp, S., Blitz, D., & van Vliet, P. (2014). Factor investing: Long-only versus long-short. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2417221>
- Ilmanen, A., Israel, R., Moskowitz, T. J., Thapar, A. K., & Lee, R. (2021). How do factor premia vary over time? A century of evidence. *A Century of Evidence (February 18, 2021)*.
- Jacobs, B. I., & Levy, K. N. (1996). 20 myths about long-short. *Financial Analysts Journal*, 52(5), 81-85.
- Jacobs, B. I., Levy, K. N., & Starer, D. (1993). Long/short equity investing. *Journal of Portfolio Management*, 20(1), 52.
- Jacobs, B. I., Levy, K. N., & Starer, D. (1999). Long-short portfolio management: An integrated approach. *Journal of Portfolio Management*, 25, 23-32.
- Jacobsen, R. (1988). The persistence of abnormal returns. *Strategic Management Journal*, 9(5), 415-430.
- 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.
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance*, 23(2), 389-416. <https://doi.org/10.2307/2325404>
- Jiang, J., Kelly, B., & Xiu, D. (2023). (Re-) Imag(in)ing price trends. *The Journal of Finance*, 78(6), 3193-3249. <https://doi.org/10.1111/jofi.13268>
- Jin, W., Li, Z. J., Wei, L. S., & Zhen, H. (2000). The improvements of BP neural network learning algorithm. In *WCC 2000-ICSP 2000. 2000 5th International Conference on Signal Processing Proceedings. 16th World Computer Congress 2000* (Vol. 3, pp. 1647-1649). IEEE.
- Li, Y., & Ma, W. (2010, October). Applications of artificial neural networks in financial economics: A

- survey. In *2010 International Symposium on Computational Intelligence and Design* (pp. 211-214). IEEE.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1), 13-37.
- Litterman, B. (2003). *Modern investment management: An equilibrium approach* (Vol. 190). John Wiley & Sons.
- Liu, S., Zhang, C., & Ma, J. (2017). CNN-LSTM neural network model for quantitative strategy analysis in stock markets. In S. Omatu et al. (Eds.), *New advances in intelligent decision technologies* (pp. 198-206). Springer. [https://doi.org/10.1007/978-3-319-70096-0\\_21](https://doi.org/10.1007/978-3-319-70096-0_21)
- Ma, L. (2020). *Quantitative investing: From theory to industry*. Springer Nature.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59-82.
- Malkiel, B. G. (2021). *A random walk down Wall Street: The time-tested strategy for successful investing*. W.W. Norton & Company Ltd.
- Margenot, M., Wassermann, G., Hall, J. C., & Granizo-Mackenzie, D. (n.d.). *Factor Analysis with Alphalens*. Quant Rocket.
- Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance*, 7(1), 71-91.
- McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1), 5-32.
- Michaud, R. O. (1993). Are long-short equity strategies superior? *Financial Analysts Journal*, 49(6), 44-49.
- Naseer, M., & Bin Tariq, D. Y. (2015). The efficient market hypothesis: A critical review of the literature. *The IUP Journal of Financial Risk Management*, 12(4), 48-63.
- Osborne, M. F. (1959). Brownian motion in the stock market. *Operations Research*, 7(2), 145-173.
- Qian, E., & Hua, R. (2006). Active risk and information ratio. In *The World Of Risk Management* (pp. 151-167).
- Reimers, N., & Gurevych, I. (2017). Optimal hyperparameters for deep LSTM-networks for sequence labeling tasks. *arXiv preprint arXiv:1707.06799*.
- Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, 11(3), 9-16.
- Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica: Journal of the Econometric Society*, 431-449.
- Samuelson, P. A. (2016). Proof that properly anticipated prices fluctuate randomly. In *The World Scientific Handbook of Futures Markets* (pp. 25-38).
- Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, 90, 106181.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk.

- Journal of Finance*, 19(3), 425-442.
- Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management*, 18(2), 7.
- Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1394-1401). IEEE.
- Stattman, D. (1980). Book values and stock returns. *The Chicago MBA: A Journal of Selected Papers*, 4(1), 25-45.
- Țițan, A. G. (2015). The efficient market hypothesis: Review of specialized literature and empirical research. *Procedia Economics and Finance*, 32, 442-449.
- Wen, Y., Lin, P., & Nie, X. (2020). Research of stock price prediction based on PCA-LSTM model. In *IOP Conference Series: Materials Science and Engineering* (Vol. 790, No. 1, p. 012109). IOP Publishing.
- Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.
- Yan, H., & Ouyang, H. (2018). Financial time series prediction based on deep learning. *Wireless Personal Communications*, 102(2), 683–700. <https://doi.org/10.1007/s11277-017-5086-2>
- Yang, S. (2021). A novel study on deep learning framework to predict and analyze the financial time series information. *Future Generation Computer Systems*, 125, 812-819. <https://doi.org/10.1016/j.future.2021.07.017>
- Yin, C., Perchet, R., & Soupé, F. (2021). A practical guide to robust portfolio optimization. *Quantitative Finance*, 21(6), 911-928.
- Zhang, F., Wang, X., & Cao, H. (2021). Turnover-Adjusted Information Ratio. *arXiv preprint arXiv:2105.10306*.
- Zhang, Y., Liao, Q. V., & Bellamy, R. K. (2020, January). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 295-305).

# APPENDIX A Data

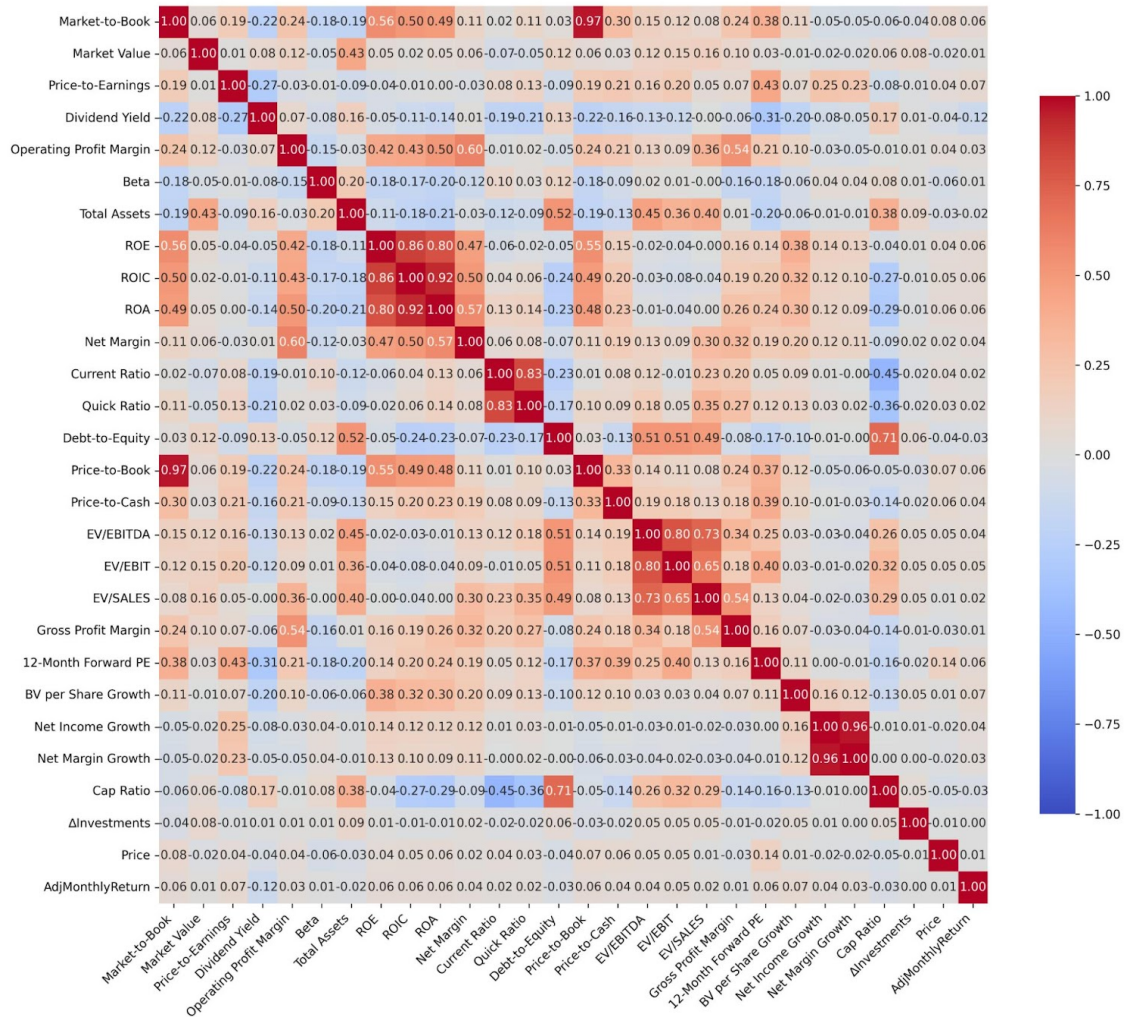


Figure 1A. Correlation matrix of the fundamental data.

Table 1A. Skewness and kurtosis p-values of the fundamental & technical data.

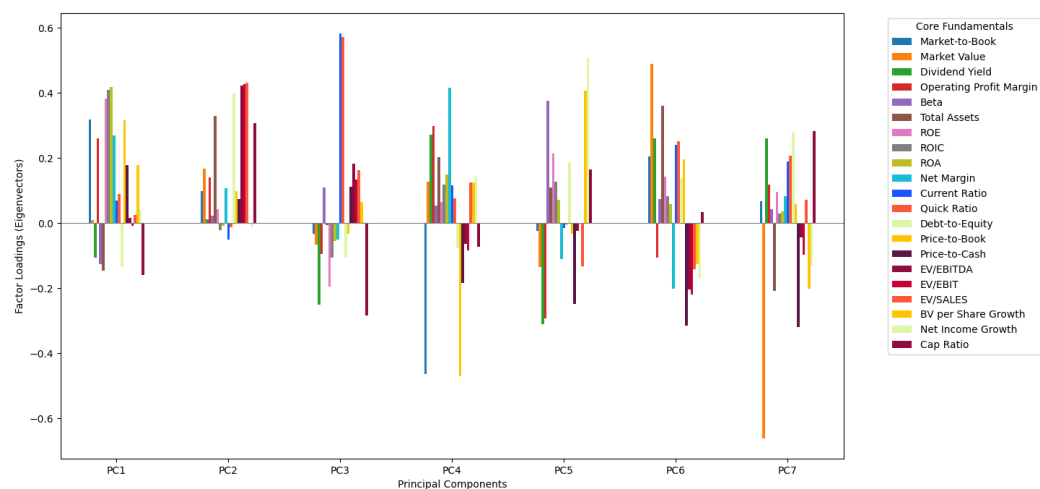
| Variable                | N     | p-value of Skewness | p-value of Kurtosis |
|-------------------------|-------|---------------------|---------------------|
| Market-to-Book          | 77892 | 0.00                | 0.00                |
| Market Value            | 80291 | 0.00                | 0.00                |
| Price-to-Earnings       | 73543 | 0.00                | 0.00                |
| Dividend Yield          | 80290 | 0.00                | 0.00                |
| Operating Profit Margin | 83517 | 0.00                | 0.00                |
| Beta                    | 79953 | 0.00                | 0.00                |
| Total Assets            | 83685 | 0.00                | 0.00                |
| ROE                     | 81817 | 0.00                | 0.00                |

|                     |         |      |      |
|---------------------|---------|------|------|
| ROIC                | 82653   | 0.00 | 0.00 |
| ROA                 | 81942   | 0.00 | 0.00 |
| Net Margin          | 83541   | 0.00 | 0.00 |
| Current Ratio       | 66669   | 0.00 | 0.00 |
| Quick Ratio         | 66741   | 0.00 | 0.00 |
| Debt-to-Equity      | 83625   | 0.00 | 0.00 |
| Price-to-Book       | 78573   | 0.00 | 0.00 |
| Price-to-Cash       | 78573   | 0.00 | 0.00 |
| EV/EBITDA           | 74127   | 0.00 | 0.00 |
| EV/EBIT             | 77978   | 0.00 | 0.00 |
| EV/SALES            | 78397   | 0.00 | 0.00 |
| Gross Profit Margin | 70046   | 0.00 | 0.00 |
| 12-Month Forward PE | 79356   | 0.00 | 0.00 |
| BV per Share Growth | 81385   | 0.00 | 0.00 |
| Net Income Growth   | 69204   | 0.00 | 0.00 |
| Net Margin Growth   | 69097   | 0.00 | 0.00 |
| Cap Ratio           | 83649   | 0.00 | 0.00 |
| ΔInvestments        | 83344   | 0.00 | 0.00 |
| Monthly Return      | 80291   | 0.71 | 0.00 |
| Open                | 1694840 | 0.00 | 0.00 |
| High                | 1694840 | 0.00 | 0.00 |
| Low                 | 1694840 | 0.00 | 0.00 |
| Close               | 1694840 | 0.00 | 0.00 |
| Adjusted Close      | 1694840 | 0.00 | 0.00 |
| Volume              | 1694840 | 0.00 | 0.00 |

---

Table 1A provides the p-values for the skewness and kurtosis for each respective fundamental and technical data variable. Market value, total assets, and change in investment are provided in billions. Other fundamental data is provided in ratios, excluding margins and growth metrics which are provided in percentages.

## APPENDIX B PCA Analysis



**Figure 1B.** The factor loadings of the core fundamentals across 7 principal components.

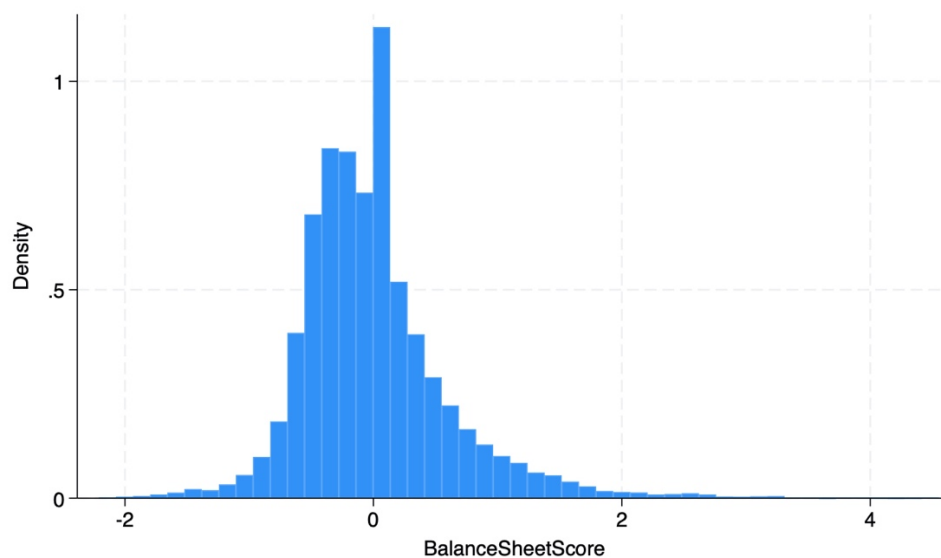
Figure 1B illustrates the eigenvectors of given core fundamentals, that were found to be both valid and stable over the in-sample period, in the PCA analysis across the first 7 principal components. The principal component is abbreviated as PC on the x-axis, and the suffix numbers indicate the number of the principal component.

**Table 1B.** Principal components of core fundamentals.

| # observations |            | 90,100     |            |            |  |
|----------------|------------|------------|------------|------------|--|
| Component      | Eigenvalue | Difference | Proportion | Cumulative |  |
| PC1            | 4.587      | 1.096      | 0.218      | 0.218      |  |
| PC2            | 3.491      | 1.319      | 0.166      | 0.385      |  |
| PC3            | 2.173      | 0.461      | 0.103      | 0.488      |  |
| PC4            | 1.711      | 0.454      | 0.082      | 0.570      |  |
| PC5            | 1.257      | 0.066      | 0.060      | 0.629      |  |
| PC6            | 1.191      | 0.158      | 0.057      | 0.686      |  |
| PC7            | 1.033      | 0.106      | 0.049      | 0.735      |  |
| PC8            | 0.927      | 0.059      | 0.044      | 0.780      |  |
| PC9            | 0.868      | 0.147      | 0.041      | 0.821      |  |
| PC10           | 0.721      | 0.086      | 0.034      | 0.855      |  |
| PC11           | 0.635      | 0.088      | 0.030      | 0.885      |  |
| PC12           | 0.548      | 0.188      | 0.026      | 0.912      |  |
| PC13           | 0.359      | 0.010      | 0.017      | 0.929      |  |
| PC14           | 0.350      | 0.089      | 0.017      | 0.945      |  |
| PC15           | 0.261      | 0.012      | 0.012      | 0.958      |  |

|      |       |       |       |       |
|------|-------|-------|-------|-------|
| PC16 | 0.249 | 0.020 | 0.012 | 0.970 |
| PC17 | 0.229 | 0.063 | 0.011 | 0.981 |
| PC18 | 0.166 | 0.019 | 0.008 | 0.988 |
| PC19 | 0.147 | 0.083 | 0.007 | 0.995 |
| PC20 | 0.064 | 0.032 | 0.003 | 0.999 |
| PC21 | 0.032 | .     | 0.002 | 1.000 |

Table 1B displays the detailed results of PCA analysis of core fundamental variables, that were found to be both valid and stable over the in-sample period. # observations indicate the number of observations being 90,100 for each core fundamental variable. This number can be higher than what has been presented in the summary statistics due to the median imputation as PCA analysis is sensitive to unbalanced data. The median imputation is preferred over the mean imputation due to each core fundamental variable showing a skewed distribution. The eigenvalues column simply informs on the magnitude of variance contained within one principal component for the core fundamentals dataset. The difference column indicates the difference of eigenvalues between two consecutive principal components. The proportion column indicates the proportion of total variance explained per principal component in the core fundamentals dataset. The cumulative column indicates the cumulative total variance of the core fundamentals dataset explained up to the given principal component.



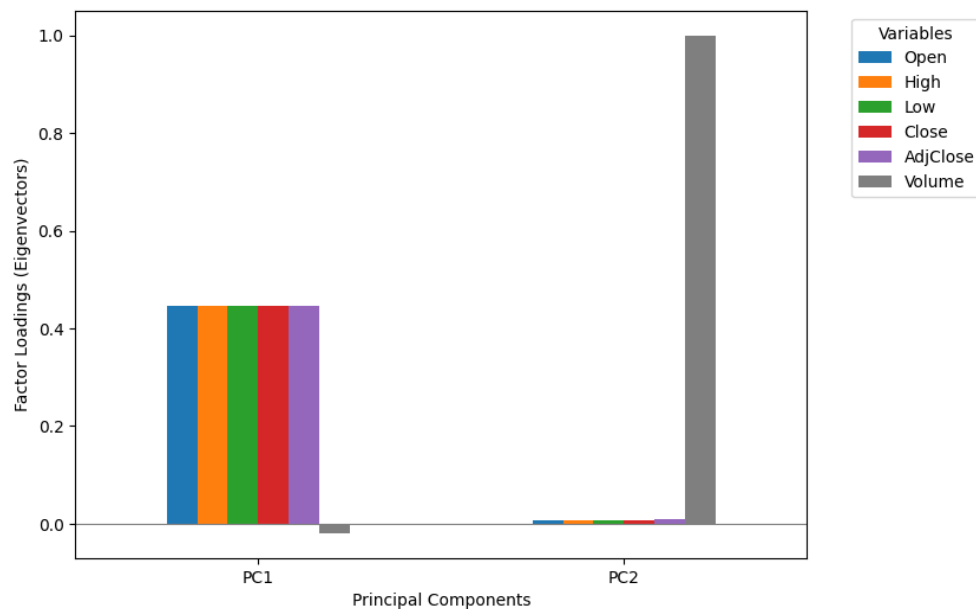
**Figure 2B.** The histogram of the balance sheet score index.

Figure 2B illustrates the density of balance sheet score index distribution, using 21 core fundamentals variables and calculated based on its PCA analysis that consists of 7 principal components. The x-axis is the balance sheet score index which is on a monthly basis on a security level.

**Table 2B.** Pairwise correlations of OHLCAV variables.

| Variables | Open   | High   | Low    | Close  | AdjClose | Volume |
|-----------|--------|--------|--------|--------|----------|--------|
| Open      | 1.000  |        |        |        |          |        |
| High      | 1.000  | 1.000  |        |        |          |        |
| Low       | 1.000  | 1.000  | 1.000  |        |          |        |
| Close     | 1.000  | 1.000  | 1.000  | 1.000  |          |        |
| AdjClose  | 0.982  | 0.982  | 0.982  | 0.982  | 1.000    |        |
| Volume    | -0.067 | -0.066 | -0.067 | -0.067 | -0.078   | 1.000  |

Table 2B provides the correlation between each pair of OHLCAV variables including – *only during the in-sample period.*



**Figure 3B.** The factor loadings of the OHLCAV variables across 2 principal components.

Figure 3B illustrates the eigenvectors of given OHLCAV variables during the in-sample period in the PCA analysis across the first 2 principal components. The principal component is abbreviated as PC on the x-axis, and the suffix numbers indicate the number of the principal component.

**Table 3B.** Principal components of OHLCAV variables.

| # observations |            | 1,694,840  |            |            |
|----------------|------------|------------|------------|------------|
| Component      | Eigenvalue | Difference | Proportion | Cumulative |
| PC1            | 4.997      | 3.998      | 0.833      | 0.833      |
| PC2            | 0.999      | 0.994      | 0.166      | 0.999      |
| PC3            | 0.005      | 0.005      | 0.001      | 1.000      |
| PC4            | 0.000      | 0.000      | 0.000      | 1.000      |
| PC5            | 0.000      | 0.000      | 0.000      | 1.000      |
| PC6            | 0.000      |            | 0.000      | 1.000      |

Table 3B displays the detailed results of PCA analysis of OHLCAV variables over the in-sample period. # observations indicate the number of observations being 1,694,840 for each core OHLCAV variable. The eigenvalues column simply informs on the magnitude of variance contained within one principal component for the OHLCAV dataset. The difference column indicates the difference of eigenvalues between two consecutive principal components. The proportion column indicates the proportion of total variance explained per principal component in the OHLCAV dataset. The cumulative column indicates the cumulative total variance of the OHLCAV dataset explained up to the given principal component.

**Table 4B.** Pairwise correlations of Actual & Predicted Price Index.

| Variables             | Actual Price Index | Predicted Price Index |
|-----------------------|--------------------|-----------------------|
| Actual Price Index    | 1.000              |                       |
| Predicted Price Index | 0.999              | 1.000                 |

Table 4B provides the correlation between the daily actual price index and the forecasted daily price index which is derived from the LSTM model - *only during the out-of-sample period*.

## APPENDIX C Alpha Regressions

**Table 1C.** OLS time-series regressions of hybrid portfolio strategy based on market benchmarks.

| Benchmark | (1)<br>Coef. (SE) | (2)<br>Coef. (SE) | (3)<br>Coef. (SE) | (4)<br>Coef. (SE)   | (5)<br>Coef. (SE)  | (6)<br>Coef. (SE)   | (7)<br>Coef. (SE) | (8)<br>Coef. (SE) |
|-----------|-------------------|-------------------|-------------------|---------------------|--------------------|---------------------|-------------------|-------------------|
| Index     | -.153<br>(.167)   |                   |                   |                     |                    |                     |                   |                   |
| Mkt-Rf    |                   | -.023<br>(.138)   |                   |                     |                    |                     |                   | -.084<br>(.124)   |
| SMB       |                   |                   | .842*<br>(.463)   |                     |                    |                     |                   | .677<br>(.513)    |
| HML       |                   |                   |                   | -1.182***<br>(.231) |                    |                     |                   | -.786<br>(.580)   |
| RMW       |                   |                   |                   |                     | 2.189***<br>(.267) |                     |                   | .457<br>(.690)    |
| CMA       |                   |                   |                   |                     |                    | -2.155***<br>(.342) |                   | -.527<br>(.659)   |
| WML       |                   |                   |                   |                     |                    |                     | .773***<br>(.256) |                   |
| Alpha     | .029***<br>(.009) | .028***<br>(.009) | .030***<br>(.009) | .029***<br>(.007)   | .026***<br>(.008)  | .022***<br>(.007)   | .026***<br>(.008) | .029***<br>(.009) |
| F-test    | 0.836             | 0.029             | 3.311             | 26.207              | 67.062             | 39.731              | 9.146             | 12.535            |
| Prob > F  | 0.365             | 0.865             | 0.074             | 0.000               | 0.000              | 0.000               | 0.004             | 0.000             |
| R-squared | 0.011             | 0.000             | 0.059             | 0.449               | 0.335              | 0.394               | 0.190             | 0.514             |
| AIC       | -137.387          | -136.794          | -140.097          | -169.577            | -159.240           | -164.309            | -148.343          | -168.479          |
| BIC       | -133.372          | -132.779          | -136.083          | -165.562            | -155.226           | -160.294            | -144.328          | -156.435          |
| N         | 55                | 55                | 55                | 55                  | 55                 | 55                  | 55                | 55                |

Table 1C provides the Ordinary Least Squares (OLS) time-series regressions of monthly rebalanced equal-weighted long-short hybrid portfolio strategy across different market benchmarks. The index represents for the S&P Europe 350 Index; Mkt-RF represents the market risk premium factor; SMB represents the size factor; HML represents the value factor; RMW represents the profitability factor; CMA represents the investment factor; WML represents the momentum factor; 5-FF represents the five-factor Fama-French model, which is a combination of Mkt-RF, SMB, HML, RMW, and CMA factors; Portfolio represents the hybrid portfolio strategy. All returns are on a monthly basis. All factor returns are retrieved from the Kenneth R. French data library. The index returns are retrieved from the FactSet database. \*, \*\*, and \*\*\* indicate significances at %10, %5, and %1 levels, respectively. Alpha metrics are calculated as seen in Equations (12). Coef stands for the respective monthly coefficient, and SE stands for the respective monthly robust standard error, which is provided in parentheses.

## APPENDIX D Maximum Drawdowns

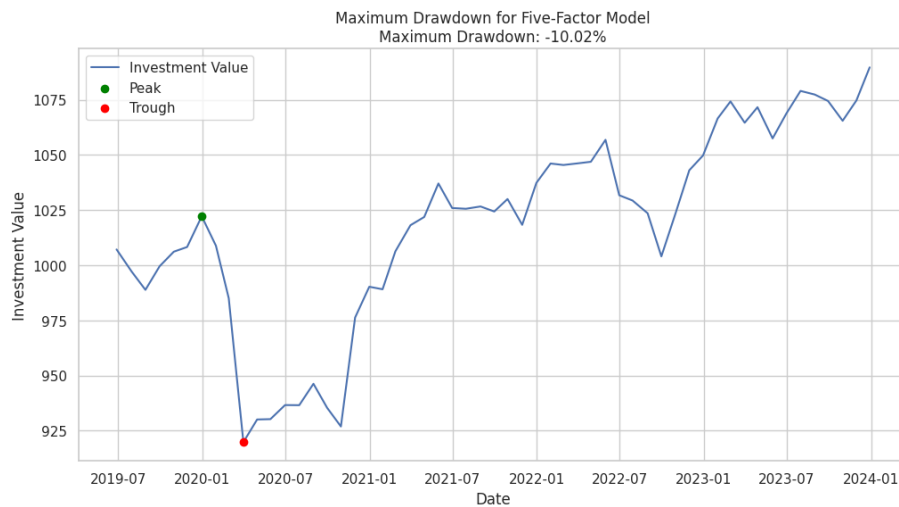


Figure 1D. Maximum drawdown of 5-factor Fama-French model (European market).

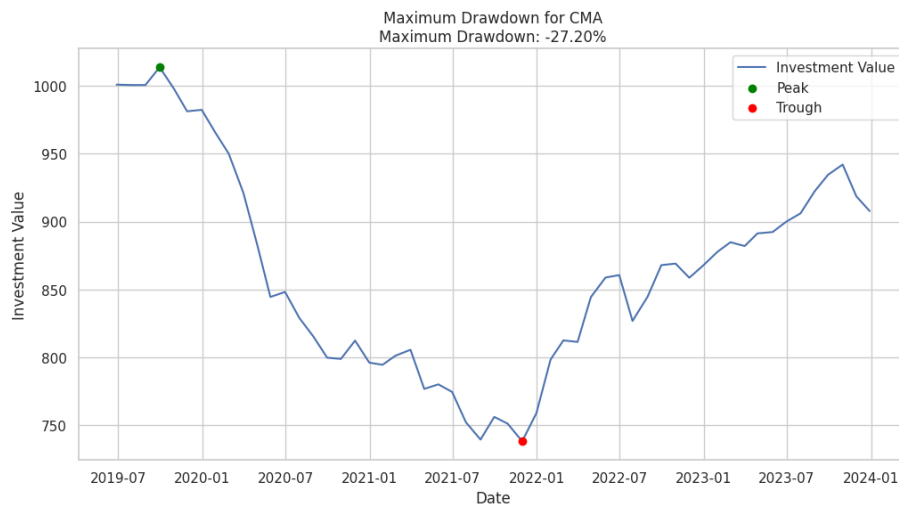


Figure 2D. Maximum drawdown of the European investment factor (CMA).

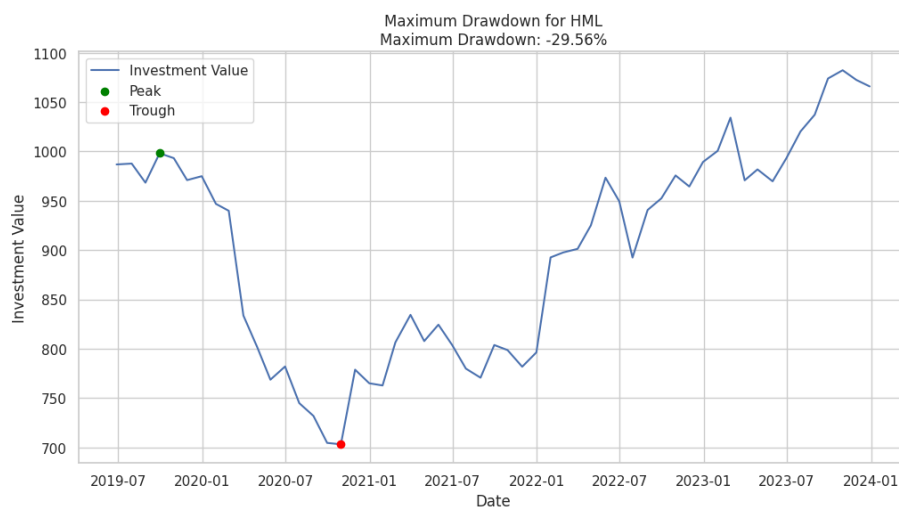
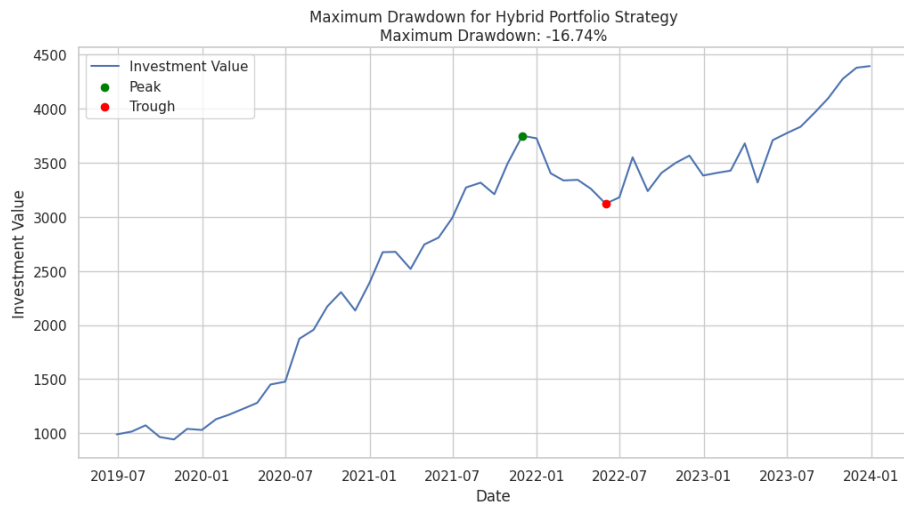
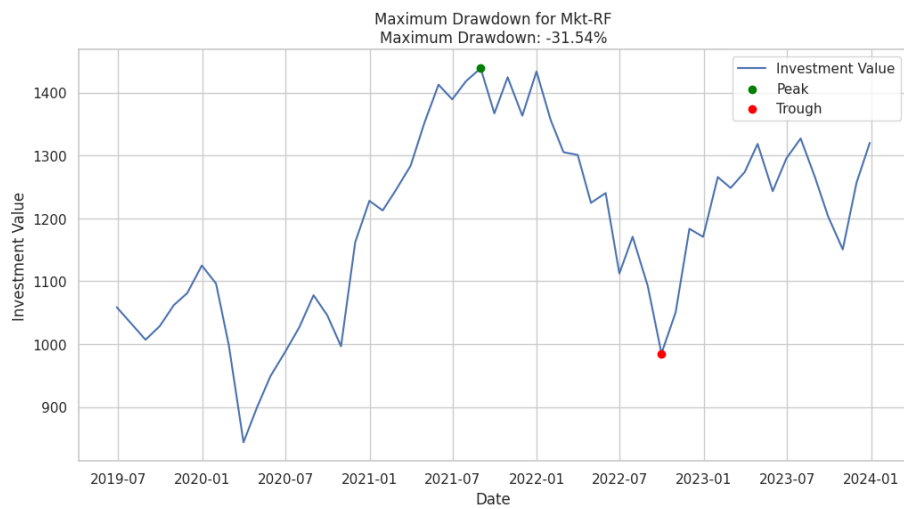


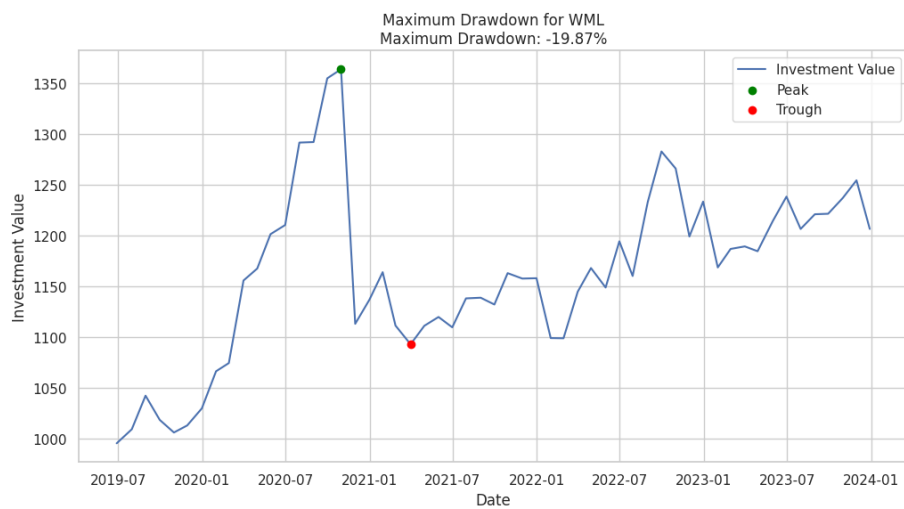
Figure 3D. Maximum drawdown of the European value factor (HML).



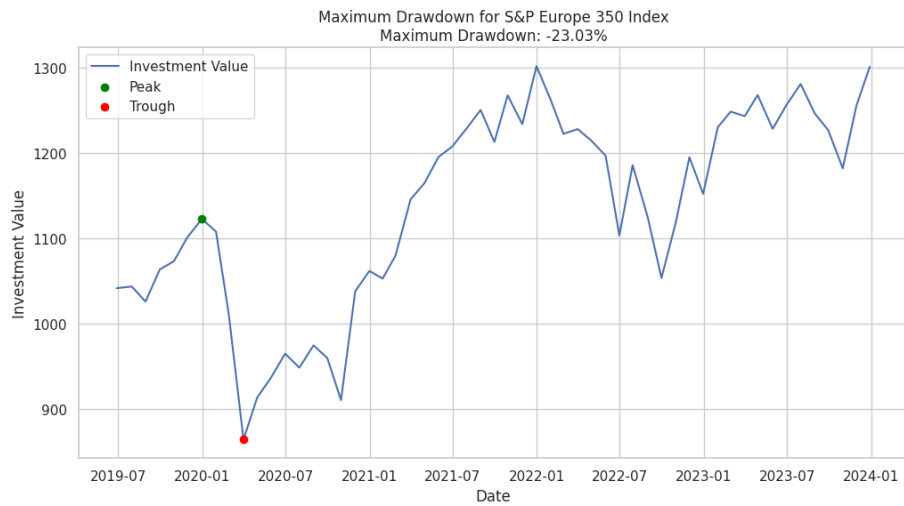
**Figure 4D.** Maximum drawdown of the hybrid portfolio strategy.



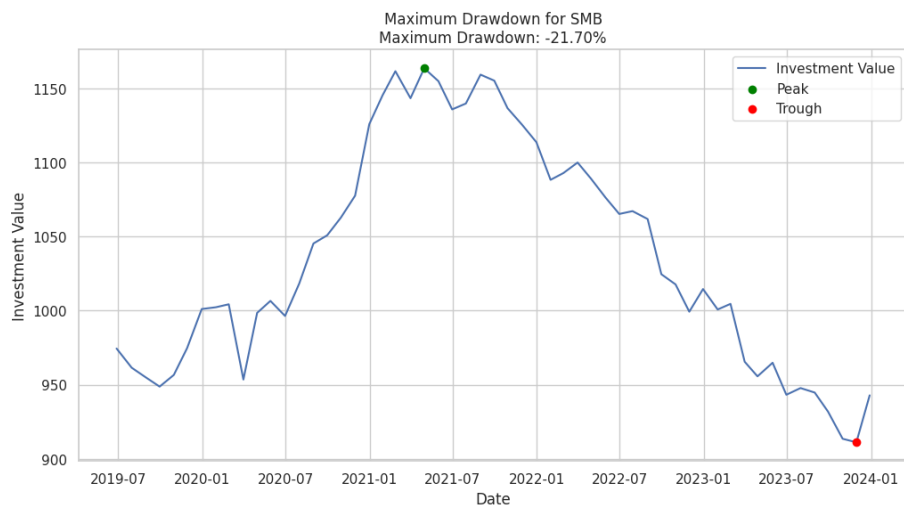
**Figure 5D.** Maximum drawdown of the European market risk premium factor (Mkt-RF).



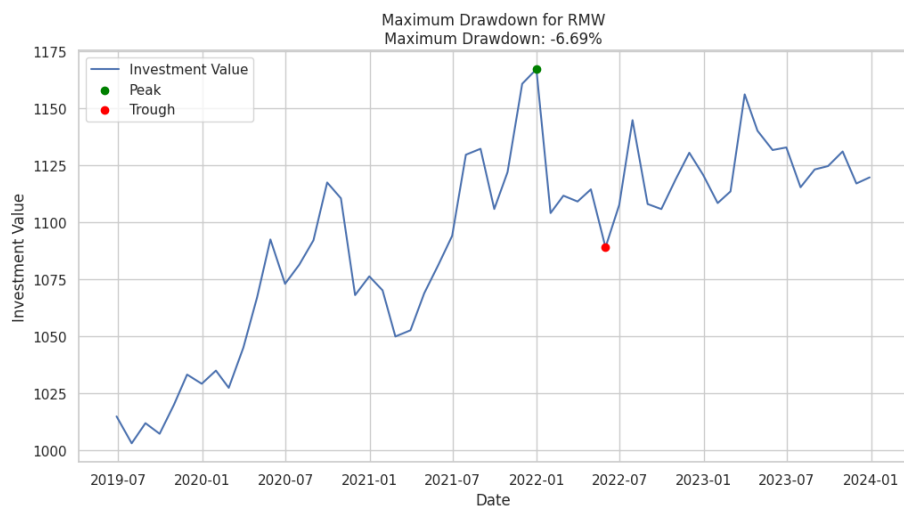
**Figure 6D.** Maximum drawdown of the European momentum factor (WML).



**Figure 7D.** Maximum drawdown of the S&P Europe 350 Index.



**Figure 8D.** Maximum drawdown of the European size factor (SMB).



**Figure 9D.** Maximum drawdown of the European profitability factor (RMW).