國立中山大學財務管理學系

碩士論文

Department of Finance

National Sun Yat-sen University

Master's Thesis

深度學習優化因子模型：以台股為例

Optimizing Factor Model with Deep Learning: A Case Study of Taiwanese Stock Market

研究生：劉維

Wei Liu

指導教授：鄭義 博士

Dr. Yih Jeng

中華民國 113 年 6 月

June 2024

論文審定書

國立中山大學研究生學位論文審定書

本校財務管理學系碩士班

研究生劉　維（學號：M104030051）所提論文

深度學習優化因子模型：以台股為例
Optimizing Factor Model with Deep Learning: A Case Study of Taiwanese Stock Market

於中華民國 113 年 6 月 12 日經本委員會審查並舉行口試，符合碩士學位論文標準。

學位考試委員簽章：

召集人 柯冠成 柯冠成　　委　員 鄭義 鄭義

委　員 翁培師 翁培師　　委　員 ______

委　員 ______　　委　員 ______

指導教授(鄭義) 鄭義 (簽名)

摘要

近年來，量化投資已成為資產管理界的重要趨勢。傳統多因子模型雖能挖掘出價值、盈餘、投資等眾多有效的量化因子，但由於採用線性迴歸等參數化方法，較難捕捉變數間的非線性關係，造成資產定價能力受到限制。而隨著深度學習等新興技術於量化研究的應用高速發展，期望能從大量金融數據中發掘更深層的投資模式與規則。

本研究之貢獻在於提供一套可用於學習共同基金投資風格的模型架構，捕捉特徵間的非線性關係，優化傳統多因子模型的資產定價能力。模型建構上參考 Feng et al. (2023)的研究框架，透過風格分析及調整模型參數，搭配台灣股市建構六大類：Momentum、Value、Profitability、Investment、Frictions、Intangible 共 51 個原始特徵，學習共同基金的投資風格。透過統計方法拆解深度學習模型的輸出結果得到重要特徵並建構因子，以加入指數增值、多空策略應用。

實證樣本期間 2007 年至 2019 年，樣本內外比約為 4:1，結果顯示：(一)透過優化參數設定及輸入(Input)，深度學習模型架構可以學習共同基金的投資風格，其資產定價能力在樣本外期間最高可達 82%。(二)使用台股市場共同基金能夠尋找到更優秀的策略，在增值指數型基金的應用中最高可達 0.99 的資訊比率，在多空策略中最高可達 1.96 的夏普值。(三)模型設定架構使得輸出(Output)包含許多動能因子特徵，複合其他特徵的動能因子能有效緩解動能崩潰。

關鍵字：風格分析、多因子模型、深度學習、量化分析、增強指數基金、投資組合管理

Abstract

Quantitative investment has become an important trend in the asset management industry in recent years. Although traditional multi-factor models can uncover many effective quantitative factors such as value, earnings, and investment, their use of parametric methods like linear regression makes it difficult to capture complex non-linear relationships, limiting their asset pricing ability. Because of this, emerging technologies such as deep learning are rapidly developing in quantitative research, hoping to discover deeper investment patterns and rules from large financial datasets.

This research contributes to providing a model framework for learning the investment styles of mutual funds, capturing non-linear relationships between features, and optimizing the asset pricing ability of traditional multi-factor models. The model construction refers to the research framework of Feng et al. (2023). Through style analysis and parameter adjustment, combined with the Taiwan stock market, six categories are constructed: Momentum, Value, Profitability, Investment, Frictions, and Intangible, with a total of 51 raw features, to learn the investment styles of mutual funds. Statistical methods are used to break down the output of the deep learning model to obtain important features and construct factors for applications in index enhancement and long-short strategies.

The empirical sample period is from 2007 to 2019, with an approximate 4:1 ratio between in-sample and out-of-sample data. The results show: (1) Through optimized parameter settings and inputs, the deep learning model framework can learn the investment styles of mutual funds, with an asset pricing ability of up to 82% in the out-of-sample period. (2) Using Taiwan stock market mutual funds, superior strategies can be found, with a maximum information ratio of 0.99 for enhanced index fund applications and a maximum Sharpe ratio of 1.96 for long-short strategies. (3) The model framework incorporates many momentum factor features in the output, and the composite momentum factor with other features effectively mitigates momentum crashes.

Keywords: Multi-Factor Model, Deep Learning, Portfolio Management, Quantitative Analysis, Enhanced Index Fund, Style Analysis

Content

List of Figures

List of Tables

1. Introduction

1.1 Background and Motivation

Since the world's first ETF (Exchange Traded Fund) - SPDR S&P 500 ETF (ticker: SPY) was listed in the United States in 1993, it marked the beginning of the ETF development. In light of the trend toward globalized investment, the promotion of index investing philosophy, and the establishment of a sound regulatory framework, Taiwan amended its laws in 2000 to allow the listing and issuance of ETFs in the domestic market. In 2003, Taiwan's first and currently largest ETF, the Yuanta/P-shares Taiwan Top 50 ETF (ticker: 0050-TW), was listed. In recent years, besides traditional stock index ETFs, Taiwan has also successively launched other types of ETFs, such as bond ETFs, commodity ETFs, inverse/leveraged ETFs, and short ETFs, providing investors with more diverse investment tools. Following the trend of global investment, Taiwan has also listed many ETFs tracking international indices, allowing investors to easily invest in overseas markets, such as ETFs tracking indices in the United States, China, and emerging markets. Taiwan has seen a surge in ETF investment. As of the end of January 2024, Taiwan's ETF market has reached NT$4.0179 trillion, accounting for 57.61% of the overall domestic fund market size of NT$6.9746 trillion. Moreover, the ETF market grew by NT$1.5451 trillion year-on-year, reflecting a remarkable growth rate of 62.48%.

Since the objective is to track a benchmark index, index funds have the advantage of relatively stable performance. To maintain the stable performance characteristic of index investing while still having the opportunity to outperform the benchmark index, the financial market has developed 'enhanced index funds' that combine active and passive investment styles effectively. These funds adopt index-tracking as the basic strategy but incorporate moderate active management and quantitative methods. Enhanced index funds aim to generate higher returns than the benchmark index through market timing, stock selection, or leverage.

Returns can be augmented through the deployment of numerous quantitative techniques. With the development of modern portfolio theory, a large number of multi-factor quantitative analytical applications related to asset pricing models have emerged, which are suitable for application in return enhancement mechanisms. Due to the increase in computing power and the constant introduction of new financial products in financial markets, a vast array of quantitative analytical applications have also begun to flourish.

The history of quantitative trading analysis can be traced back to the 1970s when scientists started using mathematical models and computer programs to predict the behavior of securities markets. As computer technology advanced, quantitative trading applications continued to expand and entered the mainstream in the 1990s. In recent years, quantitative analysis trading has become an important trading strategy, widely used in securities, foreign exchange, commodities, and other markets. As technology progresses and data availability increases, quantitative trading will continue to develop and become more widespread. The continuous evolution of machine learning and deep learning approaches has spurred research into discovering variable relationships. More and more index funds that combine different positions or exposures to different themes are being listed.

Although quantitative trading analysis has had decades of development history, with the rapid advancement of artificial intelligence and deep learning techniques, utilizing these technologies to discover effective pricing factors and investment strategies has become a new hot topic and challenge in the field of quantitative trading. While deep learning techniques have emerged as a new trending topic in asset pricing, their practical applications still face several limitations. Models often suffer from a lack of interpretability and are viewed as black boxes. Additionally, deep learning models typically require large amounts of data for effective training. Other common issues include overfitting concerns and the challenge of different models being suitable for different scenarios. Motivated by these considerations, this research aims to leverage deep learning model architectures to identify effective pricing factors, unraveling nonlinear relationships and interactions among factor variables to enhance traditional multifactor pricing models. Furthermore, utilizing the distinctive characteristics of the Taiwan stock market data, investment strategies tailored for the Taiwan market will be constructed based on the findings.

The key to achieving strong fund performance lies in deeply understanding the stock selection logic and the value creation mechanisms behind it. Effective value factors and investment strategies require not only quantitative analysis techniques, but also a thorough understanding of company fundamentals, industry trends, and market behavior. By combining deep learning models with fundamental analysis, we aim to grasp the core value drivers, thereby designing portfolio optimization strategies with foresight. Only by combining quantitative tools with substantive insights can we uncover true investment opportunities and continue to create value for investors.

By constructing factors from these identified characteristics and applying them to portfolio enhancement as well as long-short strategy applications, we aim to comprehensively evaluate their performance. Furthermore, we will strive to develop strategies that can deliver robust alpha generation across varying market environments and conditions. Although return generation is a pivotal objective, prudent strategy implementation necessitates a comprehensive assessment of risk exposures. Alongside evaluating the return profiles of our portfolio strategies, we employ the Barra risk model to scrutinize the underlying factor exposures and forecast the dynamic risk. This multidimensional analysis ensures that our strategies not only target attractive returns but also maintain judicious risk management, a critical consideration for practical applications.

1.2 Research Purpose

1. Departing from traditional factor models that emphasize statistical fit and typically assume linear relationships, this thesis optimizes traditional factor models through deep learning, exploring the nonlinear relationships and interactions among factors. It incorporates economically meaningful activation functions to construct models that better align with practical applications, making them more intuitive.
2. To enable the training model to be more flexibly applied and to construct a training model suitable for the Taiwan stock market, we employ the Style Analysis proposed by Sharpe (1992) to construct the target portfolio for model input.

   We then optimize the parameters of the deep learning framework based on data features, accomplishing the task of learning the fund's factor ranking.
3. To address the 'black box' criticism, this research applied statistical methods to decompose the important features from the deep learning results. We then construct the applicable factors by Principal Component Analysis (PCA) for practical applications in commoditization and investability.
4. To identify strategies applicable to the Taiwanese stock market, this study compared

the learning outcomes of various fund types across multiple dimensions, including factor effectiveness, risk analysis, and portfolio strategy application. Furthermore, it was found that due to the framework settings of the research model, the combination of different types of factors can mitigate the impact of momentum crashes.

1.3 Our Research Framework

This research comprises five chapters in total. Chapter 1 introduces the background and expected goals of the research. Chapter 2 is a literature review, starting with modern portfolio theory and exploring the development process of asset pricing. It also introduces the application of emerging technologies such as machine learning and deep learning in asset pricing models, as well as summarizing and organizing the blind spots and pitfalls in the application of these new technologies. Chapter 3 outlines the methodology, explaining the details of model planning and optimization. Related analytical applications include return models and risk models. Chapter 4 presents the empirical findings, showing the performance of investment portfolio strategies generated by different learning subjects, including an analysis of practical applications through enhanced applications and long-short strategies. Chapter 5 provides conclusions and recommendations for further research.

2. Literature Review

2.1 Modern Portfolio Theory and Factor Model

The relationship between returns and risk has always been the most important issue in portfolio management. While investors pursue higher returns, they must also carefully control their risk exposure, and there is a complicated trade-off between the two. To achieve optimal portfolio construction, modern portfolio theory has continuously developed a series of fundamental models for investment decision-making. Markowitz (1952) proposed the Modern Portfolio Theory (MPT), which used a mean-variance framework to characterize the relationship between returns and risk, emphasizing that investors can achieve better returns by diversifying risk. Based on the mean-variance theory, Sharpe (1964) and Lintner (1965) proposed the Capital Asset Pricing Model (CAPM), which explains the linear relationship between asset returns and risk, using the beta coefficient to measure the sensitivity of an asset to market risk.

However, due to the rather strict assumptions of the CAPM model, which only includes a single market risk factor, it seems insufficient in explaining the actual return performance of financial assets. Cox and Ross (1976) proposed the Arbitrage Pricing Theory (APT), which theorizes that under the condition of no arbitrage, asset returns are impacted by multiple factors, and can be used to explain the prices and return volatilities of different assets. Gradually, a significant amount of research on multi-factor models has emerged. Fama and French (1995) introduced the three-factor model, augmenting the CAPM by incorporating the Size factor (SMB) and Value factor (HML), which examine common stocks of U.S. listed companies on the NYSE and NASDAQ from 1963 to 1990, the explanatory power for expected stock returns reaches 90%. Carhart (1997) further expanded upon the Fama and French three-factor model by adding the Momentum factor, providing a more comprehensive analytical framework. Gradually, research on multi-factor models emerged in large numbers. Fama and French (1995) proposed the three-factor model, adding the Value and Size factors to the CAPM model. Carhart (1997) added the Momentum factor to the Fama and French three-factor model. During this period, many scholars devoted themselves to studying factor models in different markets and found that over time, the explanatory power of the models gradually declined. The reason was that there might still be many pricing factors yet to be discovered. Scholars began to look for explanatory models from more perspectives, such as Connor (1995) who incorporated macroeconomic factors to observe economic time series, such as inflation and interest rates, as indicators to measure the general impact on securities returns. Fama and French (2015) further augmented their original three-factor model by adding the Earnings factor (RMW) and Investment factor (CMA). After testing the model over a sample period spanning more than 50 years, their empirical findings demonstrated that the five-factor model significantly outperformed the CAPM, the three-factor model, and the four-factor model proposed by Carhart (1997). In recent years, numerous optimized models and factors have been uncovered. Hou et al. (2015) introduced the q-factor model, encompassing market, size, investment, and profitability factors. Empirical studies demonstrated that the q-factor model performs at least on par with the Fama and French (1995) three-factor and Carhart (1997) four-factor models, and even excels in capturing other significant anomalies. Moreover, Ahmed et al. (2019) conducted time-series tests, cross-sectional regressions, and multiple non-nested model comparisons on models including CAPM, Fama-French five-factor model, and Hou et al. (2015), revealing that the q-factor model proposed by Hou et al. (2015) exhibited superior overall performance.

Furthermore, Hou et al. (2021) proposed the q5 factor model, which incorporates an expected investment growth factor into the q-factor model introduced by Hou et al. (2015). While maintaining current investment and expected profitability constant, firms with higher expected investment growth rates should attain higher expected returns compared to those with lower expected growth rates. Their model demonstrated robust explanatory power, yielding an average premium of 0.84% per month over a 52-year sample period. Luo (2021) used the q5 factor model of Hou et al. (2021) to examine the return performance of various factors in the Taiwan stock market. The research found that the profitability factor and the expected investment growth factor were able to generate significantly positive returns. It also compared the pricing ability of the q5 model with the Fama-French three-factor model (1995), five-factor model (2015), and six-factor model (2018). The results showed that in explaining market anomalies, the q5 factor model had superior return explanatory power compared to other multi-factor models. In this study, I constructed raw characteristics based on Hou et al. (2020).

Modern portfolio management theory focuses on exploring the relationship between returns and risk. As research on multi-factor models has become prevalent, the idea of structured risk models has gradually emerged. Structured risk factor models explain a stock's returns using a set of common factors and an idiosyncratic factor specific to that stock, and utilize the volatility of the common factors and idiosyncratic factor to explain the volatility of the stock's returns. The advantage of structured multi-factor risk models lies in their ability to reduce problem complexity by identifying important factors. As long as the number of factors remains constant, the complexity of solving the problem does not change, even if the number of stocks in the portfolio varies. This study further observes the risk performance of the factors, primarily referencing the Barra risk model proposed by Biner et al. (2009).

In application, the Barra risk model can be used to (1) Attribute past portfolio performance. Understand the reasons behind good or poor past portfolio performance by examining which risk factor exposures were abnormal, significantly impacting past performance. (2) Analyze current portfolio risk. Analyze the overall risk level of the portfolio by looking at the current exposure situations of various risk factors in the portfolio. (3) Predict future portfolio risk. Predict potential risk situations that the portfolio may face in the next phase based on the expected changes in various risk factors according to the risk model. (4) Develop investment strategies. Develop corresponding investment strategies and portfolio adjustment plans in advance based on the analysis results of the risk model to address potential future risk situations.

2.2 Deep Learning Application in Return Forecast

Benefiting from the enhancement of computational power, the application of machine learning and deep learning in financial markets has been an enthusiastic topic for many investors. Abe and Nakayama (2018) compared deep neural networks (DNN), support vector regression (SVR), and random forests (RF) in predicting Japanese stock returns, finding that deep neural networks (DNN) exhibited the best predictive performance. Regarding the performance of the Long-Short Portfolio Strategy, DNN also demonstrated the optimal risk-return ratio. Gu et al. (2020) argued that for return prediction, the optimal model is neural networks, as their study suggested that the predictive advantage stems from the model's ability to capture nonlinear interactions overlooked by other methods. They proposed that shallow learning outperforms deep learning, which may be attributed to the relatively scarce data and low signal-to-noise ratio in asset pricing problems. Their study focused on return prediction for U.S. stocks and discovered that return reversal and momentum were the strongest predictive factors. Other market studies include Tobek and Hronec (2021) and Leung et al. (2021), who conducted research on developed markets outside of the United States. Tobek and Hronec (2021) found that while the historical performance of quantitative strategies outside the United States does not help identify winning strategies in the U.S. markets, evidence from past performance in the U.S. significantly predicts returns in international markets. Hanauer and Kalsbach (2023) also focused on predicting stock returns in emerging markets. They found that compared to traditional linear models, models considered nonlinear relationships and variable interactions were able to achieve economically and statistically significant excess returns in the out-of-sample period. Even after incorporating trading costs, short-selling constraints, and limiting the investment universe to large-cap stocks, the model still delivered substantial net profits. Notably, Leippold et al. (2022) applied neural network models to the Chinese A-share market and demonstrated robust performance even during the 2015 Chinese stock market crash.

In terms of the Taiwan stock market, Chen (2019) used an Autoencoder coupled with an MLP to predict stock returns and construct an investment portfolio, backtesting on the Taiwan stock market from 2012 to 2019 over 8 years, achieving an average annualized Sharpe ratio of 1.49, outperforming the Taiwan Weighted Index. Wu (2020) employed a

CNN combined with an MLP to construct a Smart Beta trading strategy, backtesting on the Taiwan stock market from 2007 to 2017, and obtained performance superior to both the Taiwan Weighted Index and the benchmark smart beta investment portfolio.

Zong (2020) compared the application of multiple deep learning models in individual stock return prediction, including BGSA-SVM, MLP, CNN, and LSTM. The backtesting period was from 2000 to 2018 on the FTSE100 and INDU indexes. They found that except for the SVM model, whose performance was more sensitive to input data and parameter settings, the performance of the other models did not differ significantly. We know that compared to CNN and other related RNN models, MLP has a relatively simpler structure, and the model utilizes computational resources more efficiently. Therefore, we chose the MLP model as our primary application framework.

A review of the literature shows that neural network models capable of capturing non-linear relationships and interactions have better predictive ability in various markets. For factor models, neural networks could be a suitable model. However, there are some pitfalls in applying deep learning, as summarized by Cao (2023) in their study of deep learning research in financial markets. These include (1) Overfitting, where the model performs well in backtesting due to limited sample history but fails in out-of-sample testing. (2) Lookahead bias, where data leakage across sets leads to an overly optimistic evaluation of known data. (3) Implementation gap, as deep learning backtests often come with high turnover rates, and executing long-short strategies may incur high transaction costs. (4) Explainability and performance attribution, as deep learning models are 'black boxes', making it difficult to explore feature interactions.

Feng et al. (2023) proposed a structured deep-learning framework that delineates the comprehensive mechanism underpinning the fitting of cross-sectional returns via firm characteristics by constructing risk factors. An economically motivated objective function is embedded within the model, focusing on minimizing the overall pricing error. Empirical evidence on high-dimensional features highlights robust asset pricing performance and substantial investment enhancements through the identification of essential raw features. The study also proposes statistical methods to extract important features from deep learning outputs, attempting to convert the 'black box' into a 'white box'. The research sample includes the top 3000 U.S. listed companies by market capitalization from January 1972 to December 2021. The best-performing model achieves up to 72% out-sample pricing ability from January 2012 to December 2021, including the U.S. debt ceiling crises of 2011 and 2013, COVID-19, and multiple Fed rate adjustments. This demonstrates the model's robust predictive ability in the face of financial crises. In this study, we follow Feng et al. (2023) by utilizing a deep learning model to learn latent factors, and we make some improvements to adapt the model for the Taiwanese stock market.

As an increasing number of studies on asset pricing models emerge, measuring model performance has become a new issue. Kelly et al. (2019) proposed methods with economic and statistical significance to evaluate model performance. This study follows the proposed method to measure the pricing ability of models and to examine whether the pricing ability of benchmark factor models is significantly improved after incorporating latent factors.

2.3 Style Analysis

Understanding a fund's investment style is an important consideration when making investment decisions. Fund prospectuses often explicitly state that the fund manager is limited to investing in certain types of stocks. However, fund managers may shift their investment style from favoring small-cap stocks to investing in large-cap stocks depending on different market conditions and timing. Sharpe (1992) introduced the concept of Style Analysis to analyze fund managers' style migration. Through the construction of style indices and quadratic programming to solve for the fund, the changes in a fund's investment style can be analyzed. In this study, we use Sharpe's method, combined with the construction of style indices, to apply style analysis to funds in the Taiwan stock market. We then convert it into Target Portfolio format to serve as input for the learning objective. To better align with practical investment norms, the constraints are set such that the beta exposure values resulting from the regression must fall within the range of [0, 1], and the sum of all exposure values must equal 1.

2.4 Portfolio Strategy

The father of hedge funds, Alfred Winslow Jones, founded the world's first hedge fund in 1949. Through long-short operations and risk hedging techniques, he aimed to counteract the overall market downturn risk. Its performance surpassed all mutual funds during that period. The concept of market neutral strategies gradually formed - by taking long positions in stocks expected to perform better and shorting stocks expected to perform worse, it could effectively hedge against overall market risk and focus on exploiting the relative value of individual stocks, achieving absolute returns independent of market movements. Related trading strategies such as enhanced index funds, long-short, pairs trading, etc. became widely used in actual trading scenarios. This research mainly focuses on the application of enhanced index funds and long-short strategies.

Hill and Naviwala (1999) proposed that index funds can achieve enhancement by adjusting positions in securities or derivatives. Due to limitations in the research scope, we focus solely on using securities as the tool for enhancement. Fund managers can use fundamental, technical, and chip factors as stock selection criteria to increase the target return. To better measure the performance of an enhanced index fund, we need to observe two aspects: (1) Active Risk, i.e., the tracking error of the enhanced index fund; (2) Active

Return, i.e., the excess return relative to the benchmark index. We then use the Information Ratio (IR) = Active Return/Active Risk to evaluate the fund’s performance.

According to research by Gupta et al. (1999), the highest IR can be achieved when the active risk is around 2% to 4%, while Fabozzi et al. (2002) argued that the active risk around 1.75% to 3% can achieve the optimal outcome. In this study, we measure the fund's performance with active risk close to 4%.

2.5 Momentum Crash

Jegadeesh and Titman (1993) first introduced the concept of the momentum factor, where the strong get stronger and the weak get weaker phenomenon exists in the stock market. According to research by Asness et al. (2013), the momentum factor has been empirically shown to have significant factor effectiveness and application value across multiple markets. As a strong factor in the U.S. stock market, the momentum factor is widely applied in asset allocation. However, recent research has shown that in some markets, such as the Chinese A-share market, the momentum factor has experienced long periods of failure.

Cooper et al. (2004) first pointed out that momentum profits are highly dependent on market states. Daniel and Moskowitz (2016) indicated that although momentum strategies generate strong positive average returns across multiple asset classes, during panic market states with market downturns and increased market volatility, previously underperforming assets or stocks experience abnormally strong rebounds when market conditions improve, resulting in momentum crashes.

Momentum factors may face the challenge of momentum crashes. Our research framework generates multiple latent factors with momentum characteristics, suggesting that combining various categories of factors could potentially reduce the impact of momentum crashes.

3. Research Methodology

3.1 Empirical Process

The research methodology is primarily based on the structural deep learning framework proposed by Feng et al. (2023), utilizing deep learning models to learn the feature ranking engineering of the target portfolio. While we may not have insight into the specific components derived from deep learning, we can identify significant features through a two-step Fama-MacBeth regression. Empirical investigations focusing on high-dimensional characteristics have demonstrated strong asset pricing performance and substantial investment gains, accomplished through the successful identification of the key sources of raw characteristic information. Different from Feng et al. (2023), we employ a different approach to construct the target portfolio and introduce early stopping to further mitigate the impact of overfitting. We use style analysis provided by Sharpe (1992) to analyze the investment style of funds in constructing the target portfolio, making it a more suitable learning object for the Taiwanese stock market. Moreover, the robust asset pricing performance manifested can be leveraged to formulate a multi-factor return model. This model bears practical applicability for portfolio tracking and optimization endeavors, as well as subsequent risk model analyses, echoing the approach delineated by Biner et al. (2009).

Sharpe (1992) aims to decompose the overall performance of a portfolio into various investment styles, including value, growth, small-cap, and large-cap. The concept can be applied to analyze funds in the Taiwan stock market. We apply this method to construct a target portfolio suitable for the Taiwan stock market. Details will be presented in Section 3.4.

Biner et al. (2009) use multiple style factors, such as size and momentum, to explain portfolio risk. To construct a risk model with good predictive quality, the risk model is initially built by referring to the steps of Biner et al. (2009). We take a similar model structure but utilize different inputs for risk analysis. For the practical application of factor investable, we will construct tracking and enhancing portfolios. The portfolio is also based on the FTSE TWSE Taiwan Index Series using the exact replication method to track the Taiwan 50 Index. Based on this portfolio, we enhance the portfolio by increasing the weights of stocks with better factor performance and decreasing the weights of stocks with poorer factor performance. The risk model is used to analyze the changes in exposure and composition of the two portfolios, with the expectation that investing in stocks with better factor performance increases exposure to a certain type of risk. The complete empirical process is presented in Figure 3-1.

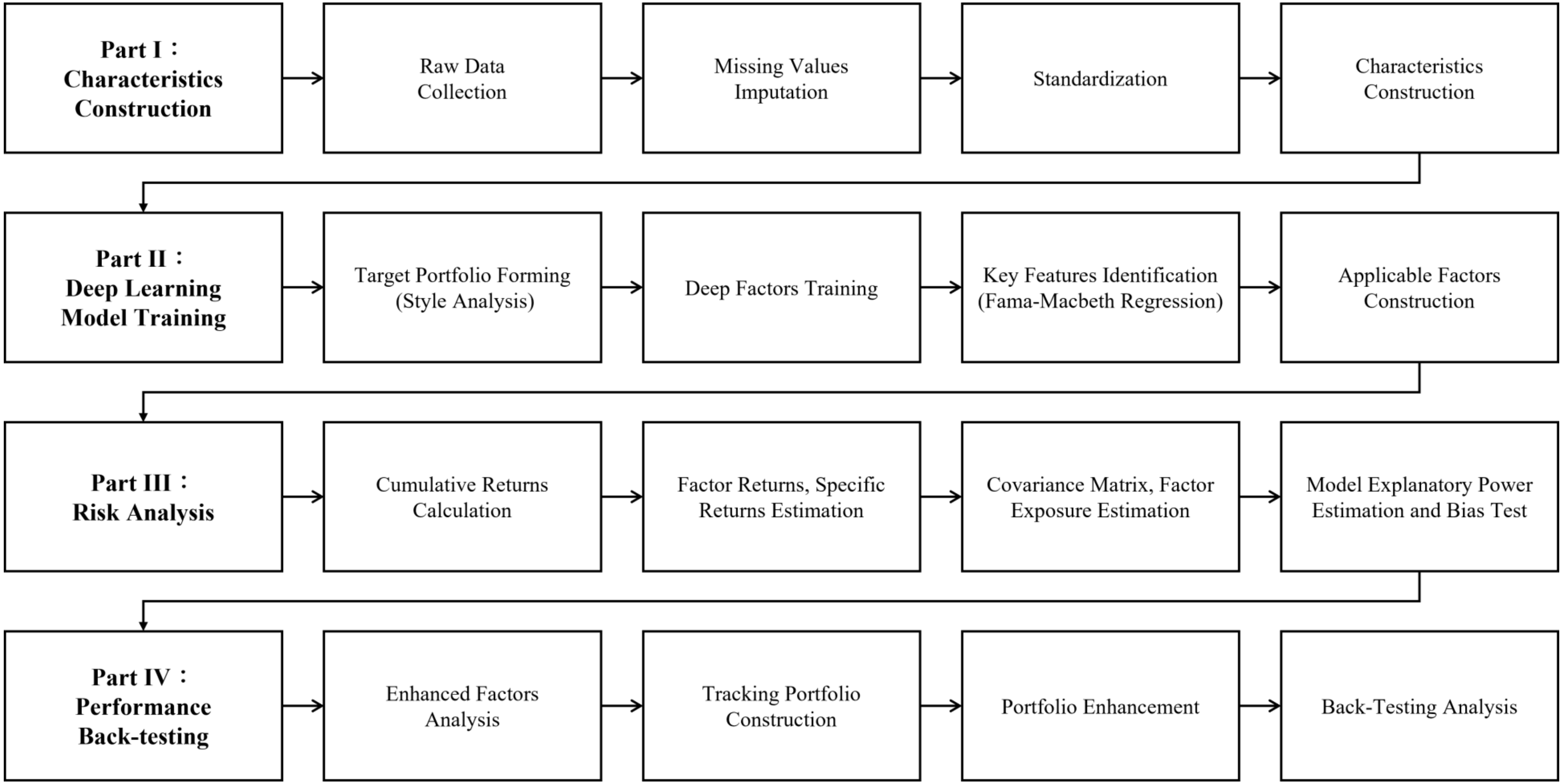

Figure 3-1 Research Framework

3.2 Data Description

The data of this research comprises individual stock trading information, individual stock financial statement data, and Taiwan stock fund data. The data are obtained from 'Taiwan Economic Journal (TEJ)'. This is for constructing characteristics and conducting style analysis on the funds. The data spanning from January 2007 to December 2019, covers various market conditions, including significant financial events such as the 2008 Financial Crisis, the 2009 Eurozone crisis, the 2014-2016 oil crisis, the ongoing US-China trade tensions since 2018, and the 2019 COVID-related financial crisis. The index data used for tracking and enhancing comes from FTSE. Table 3-1 presents the data description.

Table 3-1 Data Description

Type	Content	Period	Frequency
Stock Trading Data	Adjust Close Price Market Value Trading Volume Share Outstanding	2007/01~2020/03	Daily
Company Financial Statements Data	Income Statement Balance Sheet Statement of Cash Flows	2007/01~2019/12	Quarterly
Fund Performance	Fund Net Asset Value	2007/01~2016/12	Monthly
FTSE Index	Total Return Index Price Index	2007/01~2020/03	Daily

3.3 Definition of Characteristics

We construct characteristics based on the approach of Hou et al. (2020), which includes six categories: intangibles, value, frictions, momentum, profitability, and investment. Due to data availability constraints, a total of 51 characteristics applicable to the Taiwan stock market are eventually constructed. Construction details can be found in Appendix A.

Missing values can be filled through the backward fill method. This method fills missing values with the next available non-missing value in the series. Subsequently, the completed characteristics will be standardized according to the month.

i. Missing Value Treatment

Due to the predominant occurrence of missing values in the early section of the data period, and to enhance the volume of data points available for training, we impute the missing values. For company feature data, we employ forward filling to impute missing values. In general, if the latest company feature data is not available, we tend to use the previous feature data to fill in the missing values, assuming that the feature remains unchanged. When converting quarterly frequency data to monthly frequency data, we use backward filling to impute missing values. For example, the data for January

2015 and February 2015 would be backfilled based on the data from March 2015.

ii. Standardization

By preprocessing the input data based on its month to be standardized. We ensure the range of inputs and outputs remains consistent across every layer of the network, contributing to the overall stability of the network.

3.4 Construct Deep Factor

Based on the research conducted by Chen et al. (2024), Cong et al. (2023), and Feng et al. (2023) on the application of deep learning techniques to generate latent factors, it is believed that the augmented deep factor model can be extended to encompass the interpretable segments of both deep hidden factors and benchmark factors (e.g., CAPM or Fama-French 5 factor models). The augmented model is

$$\hat{r}_{i,t} = \beta\left(z_{i,t-1}\right)^T \mathrm{f}_t = \beta_d\left(z_{i,t-1}\right)^T \mathrm{f}_{d,t} + \beta_b\left(z_{i,t-1}\right)^T \mathrm{f}_{b,t} \tag{1}$$

$$\alpha_{i,t} = r_{i,t} - \hat{r}_{i,t} \tag{2}$$

where $\hat{r}_{i,t}$ represents the realized return predictor, which is a target portfolio serving as the learning objective. $\mathrm{f}_t = [\mathrm{f}_{d,t}, \mathrm{f}_{b,t}]$ denotes the set of deep and benchmark factors. $\beta\left(z_{i,t-1}\right)^T = [\beta_d\left(z_{i,t-1}\right)^T, \beta_b\left(z_{i,t-1}\right)^T]$ denoting the dynamic betas associated with the deep factors and benchmark factors. $\alpha_{i,t}$ is the fitting error, which evaluates the variability in both cross-sectional and time-series. The deep learning network architecture is shown as follows:

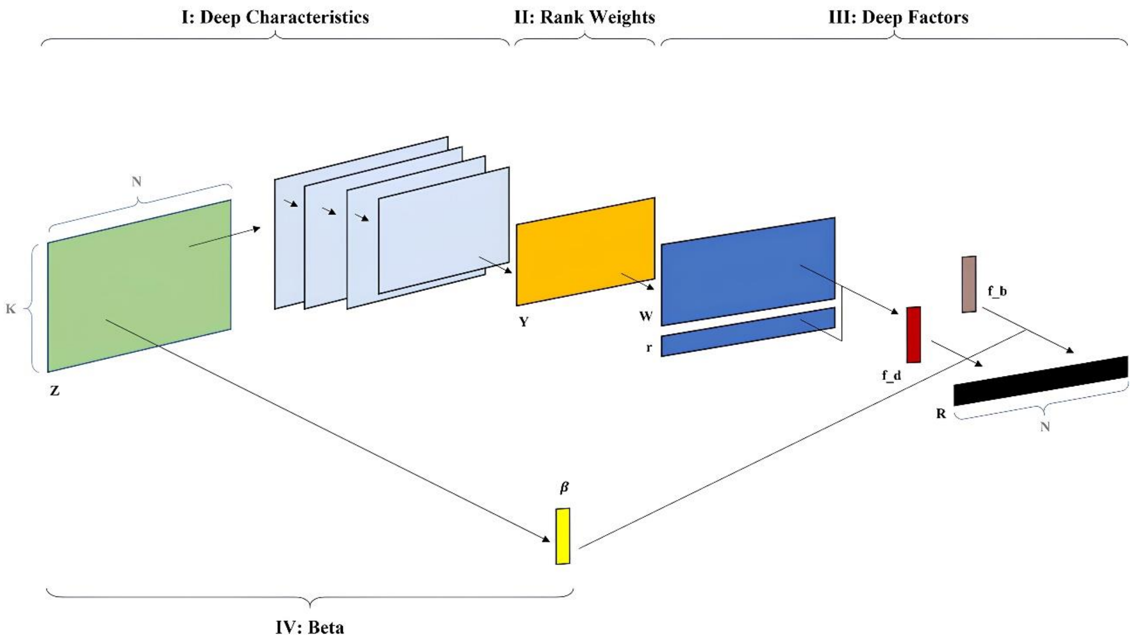

Figure 3-2 Deep Learning Architecture

To identify the interactions and non-linear relationships between features and returns, and to efficiently decompose the crucial features in deep learning, we enhanced the deep learning framework proposed by Feng et al. (2023). We commence by clarifying the symbols. A typical training observation indexed by time *t* includes the following types of data:

·$\{r_{i,t}\}_{i=1}^{N}$, excess returns of *N* individual assets;

·$\{z_{k,i,t-1}: 1 \leq k \leq K\}_{i=1}^{N}$, *K* lagged characteristics of *N* assets;

·$\{R_{b,t}\}_{b=1}^{P+1}$, a $(P+1) \times 1$ vector of excess returns on the market factor and *P* observable factors.

i. Deep Characteristics

To capture nonlinearity and interactions among the raw input features and reduce the dimension of K_0 to P_d deep characteristics. The architecture of the neural network is delineated as follows for each asset i = 1, 2,…N and l^{th} layer of the neural network (noting that the raw input corresponds to the l = 0 layer):

$$z_{i,t-1}^{(0)} = \left[z_{1,i,t-1}, \dots, z_{K,i,t-1}\right]' \tag{3}$$

$$z_{i,t-1}^{(l)} = F^{(l)}\left(A^{(l)} z_{i,t-1}^{(l-1)} + b^{(l)}\right) \tag{4}$$

for l = 1, 2…, L, where $z_{i,t-1}^{(l)}$ is the l^{th} column of the $K_l \times N$ matrix of $Z_{t-1}^{(l)}$. $F(\cdot) = (e^x - e^{-x})/(e^x + e^{-x})$, is set as a univariate activation function. The parameters to be trained in this part are deep learning weights A and biases b.

Unlike the traditional DNN setup, where every neuron is connected to all neurons in the preceding layer, in this approach, each stock is trained separately, and they do not influence each other. In this study, we fix the specifications of the neural network in the deep learning model, with [64, 16, 4] neurons in each layer, using the tanh activation function.

ii. Target Portfolio Construction

According to the empirical result in Feng et al. (2023), it can be found that using the $5 \times 5\ \mathrm{Size} - \mathrm{BM/ME}$ portfolio as the learning target shows strong asset pricing ability in both in-sample and out-of-sample in the U.S. stock market. We believe that the target portfolio applicable in different markets may vary. Therefore, we attempt to construct a similar classification portfolio structure as the learning target in the Taiwan stock market.

Sharpe (1992) introduced the method of style analysis, which involves categorizing funds based on their characteristics to analyze the active risk and active return of funds appropriately. We use this framework to construct the target portfolio.

First, this study employs the best-performing models from the main reference papers as the basis for constructing characteristic indices for the Taiwanese stock market. Also due to significant differences in returns between stocks with different value and size characteristics in the Taiwan market, this study uses value and size models as classification criteria to construct characteristic indices for the Taiwanese stock market.

Based on the market capitalization and price-to-book ratio of sample companies in the previous month, they are used as classification criteria for size and value characteristics. Listed companies are sorted by size, and the market values of larger companies are summed sequentially until the total market value reaches 70%. Companies with market values above this threshold are considered large-cap stocks, while the rest are small-cap stocks. Additionally, stocks with price-to-book ratios above the median are classified as growth stocks, and those below the median are classified as value stocks. According to the market capitalization and price-to-book ratio, a two-dimensional classification is performed, resulting in four characteristic indices: large-cap growth (LG), large-cap value (LV), small-cap growth (SG), and small-cap value (SV). Using these four characteristic indices as classified asset indices, we aim to explain the fund's style during the regression period. The regression equation and constraints are as follows:

$$R_{p,t} = \beta_{p1,t}F_{1,t} + \beta_{p2,t}F_{2,t} + \beta_{p3,t}F_{3,t} + \beta_{p4,t}F_{4,t} + e_{p,t} \tag{5}$$

$$s.t. \quad \sum_{j=1}^{4} \beta_{pj} = 1 \tag{6}$$

$$\beta_{pj} \geq 0, \text{for j} = 1 \sim 4$$

here R represents the return of the p^{th} fund in the t^{th} month, F represents the cap-weighted return rate of the n^{th} classified asset index in the t^{th} month, β represents the beta of the p^{th} fund to the n^{th} classified asset index, and e represents the unexplained part of the return variation of the p^{th} fund in the t^{th} month. The regression can be solved through quadratic programming.

Table 3-2 Style of TAIEX

Period	LG	LV	SG	SV	R^2
2000/01-2006/08	40.77%	34.39%	6.7%	18.13%	97.49%
2006/09-2013/04	40.24%	25.61%	0%	34.15%	97.42%
2013/05-2019/12	43.76%	19.79%	0%	36.45%	91.72%
2000/01-2019/12	42.77%	29.40%	3.56%	24.27%	96.89%

Table 3-2 displays the style of TAIEX, showing that large-cap growth stocks (LG) consistently have the highest proportion of explained returns across all periods. In contrast, the proportion of explained returns for small-cap growth stocks (SG) drops to 0% after August 2006, suggesting a potentially smaller impact of small-cap stocks on TAIEX during those periods. The R^2 values are consistently high, exceeding 90% in all periods.

Therefore, we use a similar method to construct the target portfolio with the following steps:

1. Select a fund.

2. Use style analysis to obtain the weights for each index.

3. Allocate the weights to the return sequences of the four characteristic indices for each period. This will ultimately result in an m-period × 4 matrix, which serves as the target portfolio for learning.

iii. Sorting and Nonlinear Weights

The $Z_{t-1}^{(L)}$ represents the matrix with P × N deep characteristics, i.e., the final output of the deep learning model we mentioned in section 3.4.i. Different from the stepwise sorting method to build zero-cost long-short factors (such as the Fama-French Size factor), we embrace a nonlinear strategy to approximate the weights of the long-short portfolio. We score the characteristics to reflect their contributions to the target variable and use the scoring vector as the input to the *softmax* function. This yields a probability distribution vector, serving as the weights for each feature. Ultimately, these weights are used to construct a long-short investment portfolio. Outlined as follows,

$$w_d(z_{t-1}) \equiv W_{t-1} = h(Z_{t-1}^{(L)}) \tag{7}$$

where the function h(·) employs *softmax* activation to compute portfolio weights determined by the ranking of deep characteristics. For the 1 × N vector x = $Z_{t-1}^{(L)}$, it appears in the form of,

$$h_{(x)} = \begin{bmatrix} softmax(x_1^+) \\ softmax(x_2^+) \\ \vdots \\ softmax(x_N^+) \end{bmatrix} - \begin{bmatrix} softmax(x_1^-) \\ softmax(x_2^-) \\ \vdots \\ softmax(x_N^-) \end{bmatrix} \tag{8}$$

where $x^+ = -50e^{-5x}$, $x^- = -50e^{5x}$,

$$softmax(x_i) = \frac{e^{x_i}}{\sum_{i=1}^{N} e^{x_i}} \; .$$

The distinctive feature of the function is it can categorize assets into three levels: high, medium, and low. Assets in the high category are assigned positive weights, those in the low category receive negative weights, and assets in the medium category are assigned zero weights. This allows for the implementation of a long-short strategy while ensuring that the portfolio maintains a zero-investment position. Next, we can calculate the deep factor, $f_{d,t}$, as follows,

$$f_{d,t} = W_{t-1} r_t \tag{9}$$

$$\hat{r}_{i,t} = \beta(z_{i,t-1})^T \mathrm{f}_t = \beta_d(z_{i,t-1})^T \mathrm{f}_{d,t} + \beta_b(z_{i,t-1})^T \mathrm{f}_{b,t} \tag{10}$$

Through the sorting process based on deep characteristics, the deep factor $f_{d,t}$ is calculated by performing the multiplication operation on matrices of the stock returns r_t and the corresponding weights W_{t-1}: $f_{d,t} = W_{t-1} r_t$. This process outlines how we formulate deep factors by combining N individual stocks in a traded portfolio.

The flexibility of deep learning enables the capturing of nonlinear betas, namely β_d and β_b, which are impacted by the firm characteristics $z_{i,t-1}$ for individual stocks. Following the nonlinear beta modeling framework proposed by Gu et al. (2021), we assume that betas are represented by the neural network structure utilizing the same high-dimensional feature set as that used for training deep characteristics.

$$\left[\beta_d\left(z_{i,t-1}\right)^T, \beta_b\left(z_{i,t-1}\right)^T\right] = F\left(z_{i,t-1}\right). \quad (11)$$

The beta and factor neural networks are jointly trained to minimize the loss function. In the empirical study, constant betas can be estimated through time series regression to model portfolio returns, even in cases where portfolio characteristics are unknown.

iv. Loss Function

The Loss function quantifies the disparity between the model predictions and the observed values. It serves as an indicator to assess the performance of the model, and the goal of optimizing the model is to minimize the Loss function. To enhance the pricing ability of the asset pricing model, we define the following Loss function,

$$\mathcal{L}_\lambda\left(\widehat{\Theta}\right) = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\left|r_{i,t} - \hat{r}_{i,t}\right| \quad (12)$$

where $\widehat{\Theta}$ is the set of all parameters in the neural networks. The objective of training the deep network is to obtain a simultaneous measurement of all bias and weight terms in the neural networks, generating factors $f_{d,t}$ and influencing the dynamics of $\beta(\cdot)$, accomplished through minimizing the objective function.

We use MAE as the loss function because we believe MAE better reflects the true prediction error and has smoother gradients in the context of deep learning models. This facilitates finding suitable directions for weight updates during training, potentially leading to easier convergence in the optimization process.

To prevent overfitting, we have implemented the Dropout technique proposed by Srivastava et al. (2014). This technique significantly enhances neural network performance by randomly omitting some neurons, thereby reducing overfitting. Additionally, we integrated an early stopping mechanism into our model. If there is no decrease in loss across twenty consecutive epochs, we immediately halt testing and document the model's outcomes. We also adjust the batch size for each training session according to the size of the dataset. Details of these optimizations are provided in Appendix B.

v. Model Performance

We use two types of R^2 to measure the explanatory power of the model. The first is the traditional R-squared, which represents the pure asset pricing ability.

$$Trad\ R^2 = 1 - \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} \left(r_{i,t} - \hat{r}_{i,t}\right)^2}{\sum_{i=1}^{N} \sum_{t=1}^{T} \left(r_{i,t} - r_f\right)^2} \tag{13}$$

We adopt the methodology proposed by (Kelly et al., 2019) to measures the statistical fitness of model as the second measurement, which can be used for individual stocks,

$$Total\ R^2 = 1 - \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(r_{i,t} - \hat{r}_{i,t}\right)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(r_{i,t} - f_{b,t}\right)^2} \tag{14}$$

it represents the proportion of realized return variation explained by the factor model-implied contemporaneous return, aggregated over all assets and all periods. This ratio is calculated by combining data from all assets and all periods. Our objective function directly relates to this *Total* R^2, aiming to minimize the cumulative realized pricing errors. To capture the optimal model results, we iterate through multiple rounds of deep learning model training, preserving records of the model with the highest *Total* R^2 in out-sample performance.

3.5 Interpreting Deep Characteristics

Traditionally, we are unable to discern the specific components within the output of deep learning. In Feng et al. (2023), the results of using linear and nonlinear decomposition methods to identify key features are similar. Here, we conduct the analysis using the linear decomposition method. In the application of factor models, we utilize the Fama-MacBeth two-step regression to examine the beta coefficients and significance of raw characteristics. It helps us understand whether these characteristics play a significant role within the deep characteristics. The identified important features will then be applied in value-added applications for portfolios. Due to the presence of multicollinearity in the independent variables of deep learning models, estimation bias issues can be eased through L1/L2 regularization. Additionally, deep learning can capture nonlinear relationships and interactions among variables, with multicollinearity issues not causing significant estimation bias. The Fama-Macbeth regression in this section is a linear regression model, necessitating attention to multicollinearity issues. Table A-2 in Appendix A displays the Variance Inflation Factors (VIF) for 51 raw characteristics. We will address high multicollinearity by removing variables with VIF>5, retaining 46 raw characteristics for the subsequent identification of important features.The model is as follows,

$$deepchar_{i,t} = a_t + b_{1,t} rawchar_{1,i,t} + \cdots + b_{46,t} rawchar_{46,i,t} + \varepsilon_{i,t} \tag{15}$$

to comprehensively address heteroskedasticity and autocorrelation issues in both cross-sectional and time series data, we apply Fama-MacBeth regression in conjunction with Newey-West adjustment standard error with 12 lags.

We apply equation (15) to compute the absolute values of beta coefficients corresponding to the most influential five raw characteristics. Following this, factor values are determined by weighting the original raw characteristics based on their respective beta coefficients.

Considering the predefined number of model layers, the deep learning model produces four deep characteristics. To capture the primary variability and reduce collinearity within the original factors, we employ PCA and set to retain the first principal component. The original factor data is fitted to the PCA model for transformation, resulting in a new single factor for portfolio enhancement. The details of value-added applications will be presented in Section 3.7 and Section 4.4.

3.6 Risk Model Analysis

In this research, we develop a multi-factor model for the previously constructed deep factors, based on the Barra risk model proposed by Biner et al. (2009). This model approached from the perspective of a pure factor model, is designed to assess portfolio risk and predict future risks. Our Estimation Universe (ESTU) comprises listed companies in the Taiwanese stock market, covering the period from March 2007 to December 2019. After accounting for missing values and the duration of companies' existence, a total of 594 companies are ultimately included in our sample.

The risk model categorizes factors into market risk (national market factor), thirty-three industry factors, and four deep factors. In the following sections, we will provide a comprehensive overview of the construction process of the risk model.

i. Standardization

As the model comprises various types of factors with different units or scales, the different units or scales among them may significantly impact the model's performance. The standardization process used in the Barra model differs from the typical standardization process. The advantage of using market capitalization weighting lies in its reflection of market size, as it gives greater influence to larger companies. The model is configured as follows:

$$X_{n,k,m}^{(std)} = \frac{X_{n,k,m}^{(raw)} - \mu_{k,m}}{\sigma_{k,m}} \tag{16}$$

$X_{n,k,m}^{(raw)}$ denotes the raw exposure of asset n to factor k for the m^{th} month.

$\mu_{k,m}$ denotes the cap $-$ weighted exposure of factor k for all assets in the m^{th} month.

$\sigma_{k,m}$ denotes the equal weighted standard deviation of factor k for all assets in the m^{th} month.

ii. Outliers Treatment

Outliers can have a notable impact on the performance of models, so they need to be removed. We apply winsorziation as a method. Data exceeding the range of ±3 standard deviations will be compressed. The operational method is as follows:

$$\tilde{X}_{n,k,m}^{(std)} = \begin{cases} 3 \times \left(1 - s_{(+)}\right) + X_{n,k,m}^{(std)} \times s_{(+)}; & X_{n,k,m}^{(std)} > 3 \\ X_{n,k,m}^{(std)}; & -3 < X_{n,k,m}^{(std)} < 3 \\ 3 \times \left(1 - s_{(-)}\right) + X_{n,k,m}^{(std)} \times s_{(-)}; & X_{n,k,m}^{(std)} < -3 \end{cases} \tag{17}$$

$$s_{(+)} = \max\left[0, \min\left[1, \frac{0.5}{\max\left(X_{n,k,m}^{(std)}\right) - 3}\right]\right] \tag{18}$$

Here, $\tilde{X}$ represents the exposures after winsorization, and X represents the original exposures before processing. The scale factor, $s_{(+)}$, compresses the positive tail of factor such that $\tilde{X} \leq 3.5$. The scale compression factor, $s_{(-)}$, of the negative tail is calculated analogously.

iii. Factor Return

After standardizing factor exposures and handling outliers, we use market factor, industry factors, and the deep factors generated by the deep learning model to construct a multi-factor model. The model structure is as follows:

$$r_n = f_M + \sum_{i=1}^{34} X_{i,n}^{I} f_I + \sum_{d=1}^{4} X_{d,n}^{D} f_D + u_n \tag{19}$$

Here r_n is the excess return of asset n, f_M is the market factor return, [$X_{i,n}^{I}$, $X_{d,n}^{D}$] represents the set of industry and deep factor exposure of asset n and [f_I, f_D] represents the set of industry and deep factor returns. Finally, u_n represents the idiosyncratic risk of asset n. The details of the cross-sectional solution are presented in Appendix C.

Next, we can compute the model's explanatory power by considering the ratio of the sum of squared residuals to the total sum of squared returns. It's important to note that this R^2 is calculated using a 12-month moving average.

$$Barra\ R^2 = 1 - \frac{\sum_{n=1}^{N} w_n\, u_n^2}{\sum_{n=1}^{N} w_n r_n^2} \tag{20}$$

$$w_n = \sqrt{mv_n} \tag{21}$$

iv. Factor Risk

To predict portfolio risk, we need to calculate the daily covariance matrix of factor returns. Different from the traditional direct computation of the covariance matrix of factor returns, considering that data from the distant past should have less influence and more recent data should have more influence, Biner et al. (2009) introduced the Exponentially Weighted Moving Average (EWMA) method with the concept of half-life. The details of constructing the daily covariance matrix of factor returns are provided in Appendix C.

v. Risk Forecasts and Bias Test

The risk model predicts the portfolio risk for the next month by using the portfolio weights for the current month, weighted by the exposure of the portfolio to each factor. The portfolio risk can be estimated as follows:

$$X_k^P = \sum_{n=1}^{N} h_n^P X_{nk} \tag{22}$$

$$sdev(r^P) = \left[\sum_{k,j=1}^{K} X_k^P F_{k,j} X_j^P\right]^{\frac{1}{2}} \tag{23}$$

$$F_{k,j} = cov(f_k, f_j) \tag{24}$$

Next, we use a Bias Test to scrutinize the prognostic prowess of the risk model during out-sample. This test is used to determine what proportion of risk predictions falls within the 95% confidence interval. We first calculate the z-scores as the measure of statistical significance.

$$z_{t,q} = \frac{r_{t+q}}{\sigma_t} \tag{25}$$

Here σ_t is the risk prediction at time t, r_{t+q} is the actual return at time t+q, q is set to one month in this research.

With perfect forecasts, the z-scores exhibit a standard deviation of one. Next, we calculate the bias statistic, $b_{t.T}$, is defined as the sample standard deviation of the z-scores values across the testing horizon, T:

$$b_{t.T} = (\frac{1}{T-1}\sum_{s=t-T+1}^{t}(z_{s,q} - \bar{z}_q)^2)^{1/2} \tag{26}$$

Assuming r_{t+q} follows a normal distribution, we can provide a 95% confidence interval for the bias statistic:

$$C_T = \left[1 - \sqrt{2/T}, 1 + \sqrt{2/T}\right] \tag{27}$$

With perfect forecasts and normal returns, 95% of bias statistic values within a broad group are expected to fall within the confidence interval C_T. Outliers of $z_{t,q}$ can have a significant impact on bias statistics. Hence, we adopt an alternative approach to mitigate the impact of outliers by employing winsorization on the z-scores:

$$\tilde{z}_{t,q} = \max(-3, \min(+3, z_{t,q})) \tag{28}$$

They are significantly less influenced by outliers compared to the raw bias statistics, which employ un-truncated z-scores.

3.7 Index Tracking and Enhanced Portfolio

i. Tracking Portfolio

According to the ground rules established by FTSE TWSE Taiwan Index for Taiwan 50 Index, we construct a tracking portfolio using the fully duplicated method. For ease of calculation, we make the following assumptions: (1) No consideration for market frictions, and (2) Adjusted closing prices are used as buy and sell prices.

ii. Investable Factor

In practice, we construct investable factors by analyzing the significant features (derived from the calculations in Section 3.5) using the method proposed by Kozak et al. (2020). We compose investable factors by weighting the first principal component coefficients obtained through PCA. Subsequently, we evaluate the factor's investment performance using cumulative return charts, cumulative IC, and rank-related correlation coefficients.

iii. Enhanced Portfolio

According to the observation in Section 3.7 ii regarding the measurement of factor investability, we implement a long-short strategy to maintain the total weight of the enhanced portfolio h_a at 0. The objective function and the constraints is as follows:

$$h_p = h_b + k * h_a \tag{29}$$

$$s.t. \sum h_p = 1, \sum h_b = 1, \sum h_a = 0 \tag{30}$$

Where h_p is the weight matrix of all stocks, h_b is the weight matrix of tracking portfolio, h_a is the weight matrix of the enhanced portfolio and k is the total portion of the parts we use for value addition.

For the value addition part, we first rank the factor and distribute all stocks into five groups. Assigned higher weights to the better-performing group and reduced weights from the weakening-performance group. We then adjusted the portions of the parts to be equal percentages, as shown in Figure 3-2.

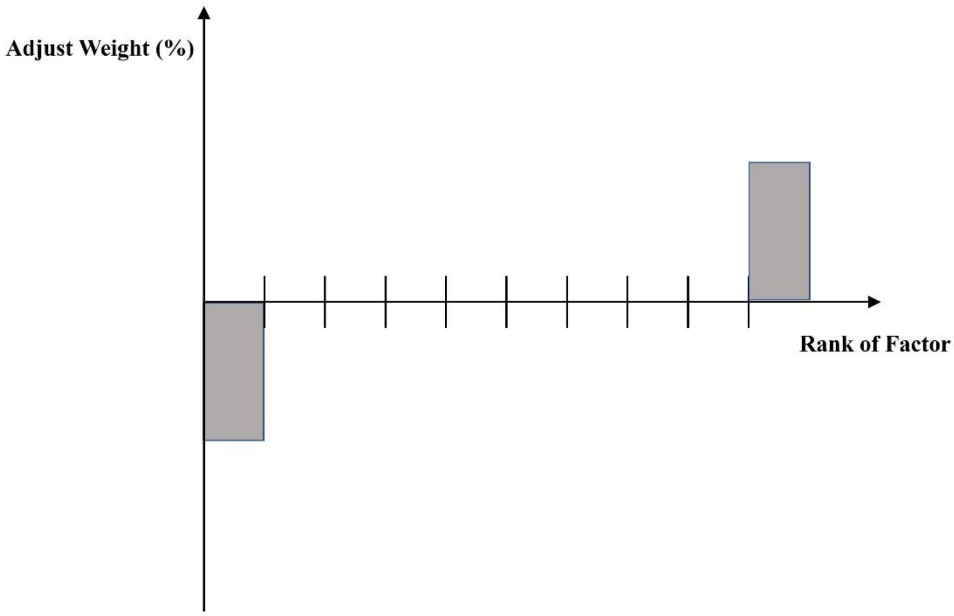

Figure 3-3 Weight Adjustment

In practice, there are constraints on shorting stocks. If the adjusted weight for an individual stock produces a negative weight after our adjustments, we set it to the minimum holding percentage for that month. We then evenly distribute the remaining portions not shorted to the other designated stocks to be shorted.

4. Empirical Results

4.1 Mutual Fund Selection and Target Portfolio Construction

Fund data is obtained from TEJ PRO. To prevent the inclusion of future data in subsequent portfolio strategy applications, the data period is set from December 2006 to December 2016. The funds included are 265 open-end equity funds investing in Taiwan stocks, denominated in both NTD and USD, with risk-return ratings of RR4 and RR5. Descriptive statistics are shown in Figure 4-1:

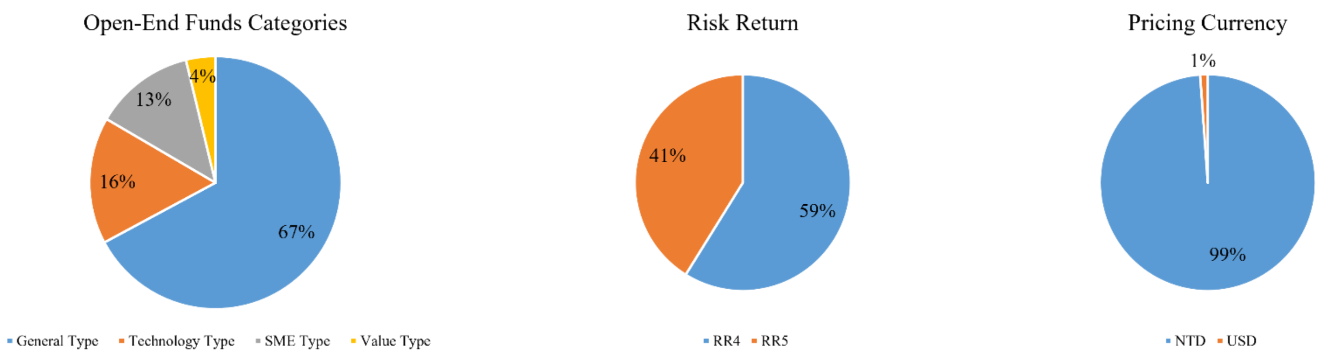

Figure 4-1 Descriptive Statistic of Mutual Fund

To ensure the robustness of our model, we created diverse types of target portfolios using different fund classification methods or indicators. We chose the following fund types for construction. The abbreviations in parentheses represent each portfolio, and subsequent empirical studies will employ these abbreviations for presentation: the top ten performers based on historical performance (TOP_10), the highest Sharpe Ratio and the best since inception during in-sample period (BEST_SHARPE/BEST_SR), open-end general type (GEN), open-end SME type (SME), open-end value type (VALUE), and

Fama 5*5 size/book-to-market (FF). The detailed descriptions of the portfolios are in Table 4-1:

Table 4-1 Descriptions of Portfolios

Portfolio	Description
TOP_10	We selected the top-performing ten funds each year based on the annual performance calculated as (NAV at the end of the year - NAV at the beginning of the year) / NAV at the beginning of the year, to serve as the targets for style analysis.
BEST_SHARPE	The calculation of the in-sample Sharpe ratio for individual stocks reveals that 'T2208Y Fuh-Hwa Digital Economy Fund' exhibited the best performance. Since its inception, the fund has achieved a return of 126%, and its Sharpe ratio stands at 0.72.
BEST_SR	Our calculation method involves comparing the NAV at the end and beginning of the sample period. The fund that demonstrated the best performance in NAV growth is the 'T0911Y UPAMC Quality Growth-A' fund, with a return rate of 126% over the period. Additionally, the fund achieved a peak return rate of 139% during the period.

GEN/SME/VALUE	Based on the fund type codes compiled by the Securities Investment Trust and Consulting Association of the R.O.C (SITCA), select the corresponding fund types and perform style analysis.
FF	The best-performing model in Feng et al. (2023). Using Size and Book-to-Market for double sorting.

We then conducted a style analysis, and the results are as follows:

Table 4-2 Style Analysis of Funds in Taiwan

Portfolio	LG	LV	SG	SV	R^2
TOP_10	35.92%	3.10%	52.16%	8.82%	78.04%
BEST_SHARPE	37.31%	10.26%	52.43%	0.00%	77.53%
BEST_SR	38.79%	0.00%	61.21%	0.00%	81.28%
GEN	36.18%	8.31%	48.61%	6.91%	77.75%
SME	25.08%	0.00%	65.18%	9.74%	78.90%
VALUE	30.46%	11.66%	41.07%	16.81%	82.71%

FF as the target portfolio was constructed based on Feng (2023) and not converted from the style analysis we conducted for funds in the Taiwan stock market.

Based on Table 4-2, we can see that when there are fewer targets measured, the R^2 tends to be higher. For BEST_SR, the index weight only includes growth stocks, TOP_10, BEST_SHARPE, and SME have relatively high index weights in growth stocks, indicating that the funds with better performance tend to invest more in growth stocks. SME small-cap stocks account for 75% of the index weight, and VALUE has a 25% index weight on value stocks, which is quite consistent with the characteristics of their fund styles. By analyzing the performance of funds within the sample through style analysis, growth stocks have a higher return explanatory power and index weight compared to value stocks. We then use these index weights to weight the time-series returns of the original characteristic indices to form the final target portfolio.

4.2 Model Asset Pricing Performance

We apply the deep learning model mentioned in Section 3.4 to take 51 Raw Characteristics as input and learn the factor ranking process for the given Target Portfolio. Optimizing the learning process through the setup of a loss function and early stopping mechanism. After splitting the dataset, we divided the samples into in-sample from March 2007 to December 2016, and out-sample from January 2017 to December 2019. To obtain better results, we ran 10 trials for each model and recorded the results from the trial with the best out-sample performance. Recording the best out-sample performance has the advantage of validating the model's ability to generalize and make predictions on new data, as well as preventing overfitting. Table 4-4 shows the asset pricing performance of each Target Portfolio in in-sample and out-of-sample.

Table 4-3 Model Performance

Target Portfolio	Trad IS R^2	Trad OS R^2	Total IS R^2	Total OS R^2
TOP_10	97.36%	74.13%	88.18%	45.12%
BEST_SHARPE	97.07%	75.15%	84.46%	44.88%
BEST_SR	97.75%	77.10%	85.90%	46.01%
GEN	96.71%	78.71%	83.47%	52.27%
SME	97.69%	72.25%	92.47%	51.90%
VALUE	94.37%	82.23%	60.57%	56.12%
FF	90.66%	61.14%	78.11%	44.99%

Although there are differences in the asset pricing ability between in-sample and out-sample data, our main focus is on whether the asset pricing capability improves after introducing deep learning factors. To examine this, we additionally plot the training loss and test loss to observe whether the model exhibits overfitting. Figures 4-2 to 4-8 will present the training and test losses during the training process, helping us determine whether the model and parameter settings show signs of overfitting.

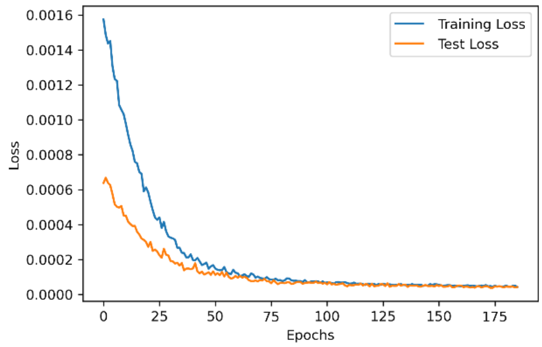

Figure 4-2 TOP_10

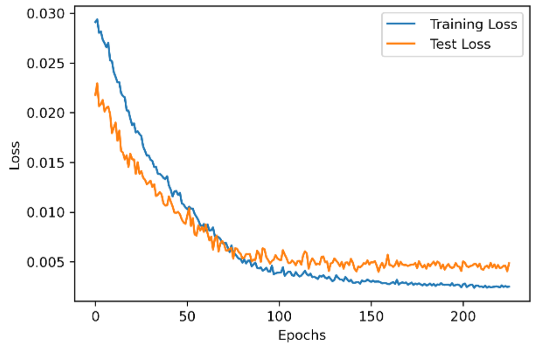

Figure 4-3 BEST_SHARPE

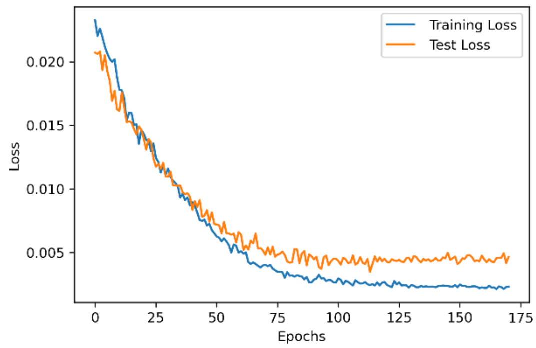

Figure 4-4 BEST_SR

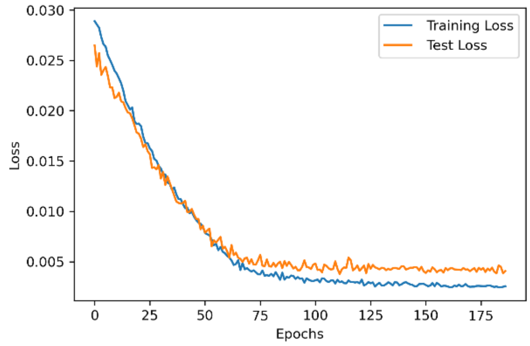

Figure 4-5 GEN

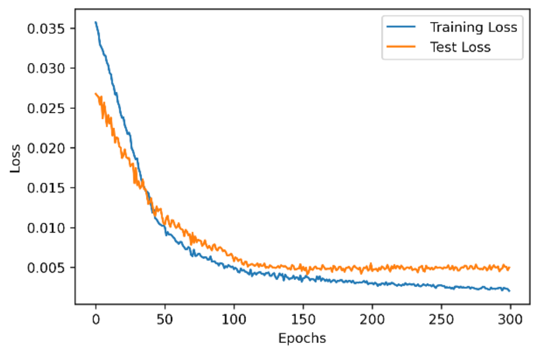

Figure 4-6 SME

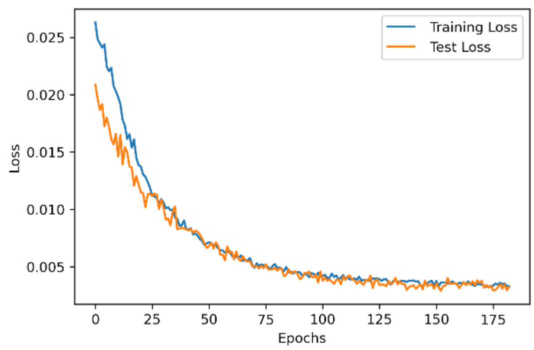

Figure 4-7 VALUE

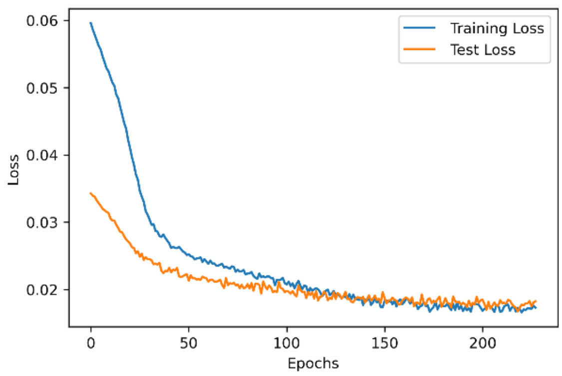

Figure 4-8 FF

In all the figures, both training and test losses decrease as the number of training epochs increases, indicating that the model's performance improves during the training process. If both training and test losses continue to decrease and the gap between them remains small, it suggests that there is no significant overfitting.

Based on model performance results, we can find that when using Trad R^2 to measure the model's asset pricing capability, most of the models exceed 70% in the out-sample, indicating that even in the presence of the COVID impact in 2019, the deep learning architecture can still effectively learn asset pricing factors. However, when we use Total R^2 to measure the model, focusing primarily on the explanatory performance of the deep factors, we observe a general decline in out-sample performance, ranging between [0.44, 0.57]. VALUE demonstrates the best performance in both R^2 measurement methods. FF, on the other hand, performs relatively poorly compared to all other models. Overall, in terms of asset pricing capability, the models exhibit good performance. Models trained using Taiwanese stock funds as the target portfolio outperform FF.

4.3 Factors and Risk Analysis

Our deep learning framework generates four deep factors. Through Fama-Macbeth regression, we identify the significant features of each deep factor, retaining the features with the top five highest beta coefficients to construct the deep factors. To gain a comprehensive understanding of the factor risk structure, in the risk model analysis section, we analyze each target portfolio using four deep factors. For future portfolio enhancement applications, we construct a single factor by weighting the four deep factors with the first principal component derived from PCA. For the convenience of illustrating subsequent empirical findings, we will name the factors constructed using PCA with their target portfolio. The table below displays the main characteristic exposures of each factor.

Table 4-4 Key Features of Deep Factors

Factor	Key Features
TOP_10	Momentum(Mom1m, Mom12m, Abr), Profitability(Pm), Intangible(Rdm)
BEST_SHARPE	Momentum(Mom1m, Nincr), Value(Sgr), Profitability(Pm), Investment(Noa)
BEST_SR	Momentum(Mom1m, Mom6m, Nincr), Frictions(Rvar_capm, Rvar_mean)
GEN	Momentum(Sue, Abr, Mom1m), Value(Cfp), Frictions(Turn)
SME	Momentum(Abr, Mom1m, Mom6m), Frictions(Maxret), Intangible(Alm)
VALUE	Momentum(Mom1m), Value(Cash, Dy), Investment(Chcsho), Frictions(Beta)
FF	Momentum(Abr), Investment(Ni, Pctacc), Frictions(Rvar_capm, Turn)

We then establish nine style factors in the Barra model to ensure that the deep factors we have learned do not exhibit excessively high correlations. The correlation coefficient matrix is shown below.

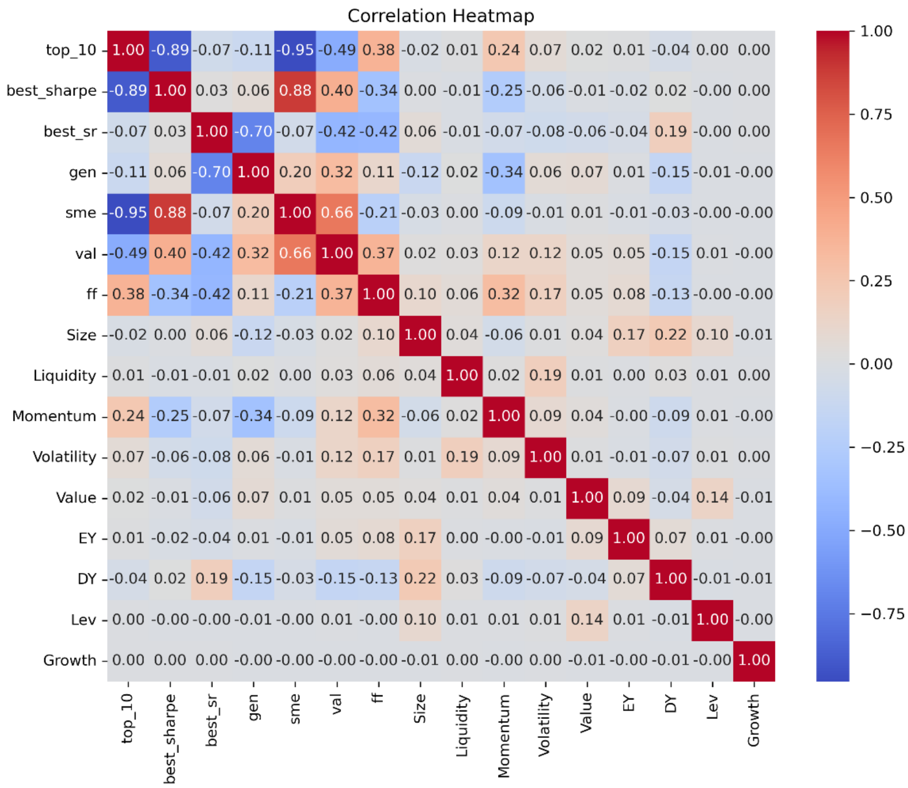

Figure 4-9 Correlation Matrix of Factors

It can be observed that, except for the relatively high correlation between GEN, FF, and momentum factor, the correlations between the other deep factors and the Barra style factors are all below 0.3. This indicates that the factors constructed through deep learning exhibit distinct factor exposures.

All the key features of our deep factors include Momentum, leading us to infer that the deep learning framework used for training may have been based on learning factor rankings of the target portfolio, combined with the Fama-French 3 factors. Besides (1) market risk, (2) the outperformance of small-cap companies relative to large-cap companies, and (3) the outperformance of high book-to-market value companies versus low book-to-market value companies, Momentum appears to be the factor that could enhance the factor model's pricing ability the most. Due to the high correlation of our factors with the traditional Momentum factor, we found that the composite factors we derived effectively mitigate momentum crashes. We then compare the three deep factors with higher correlations to the traditional momentum factor separately.

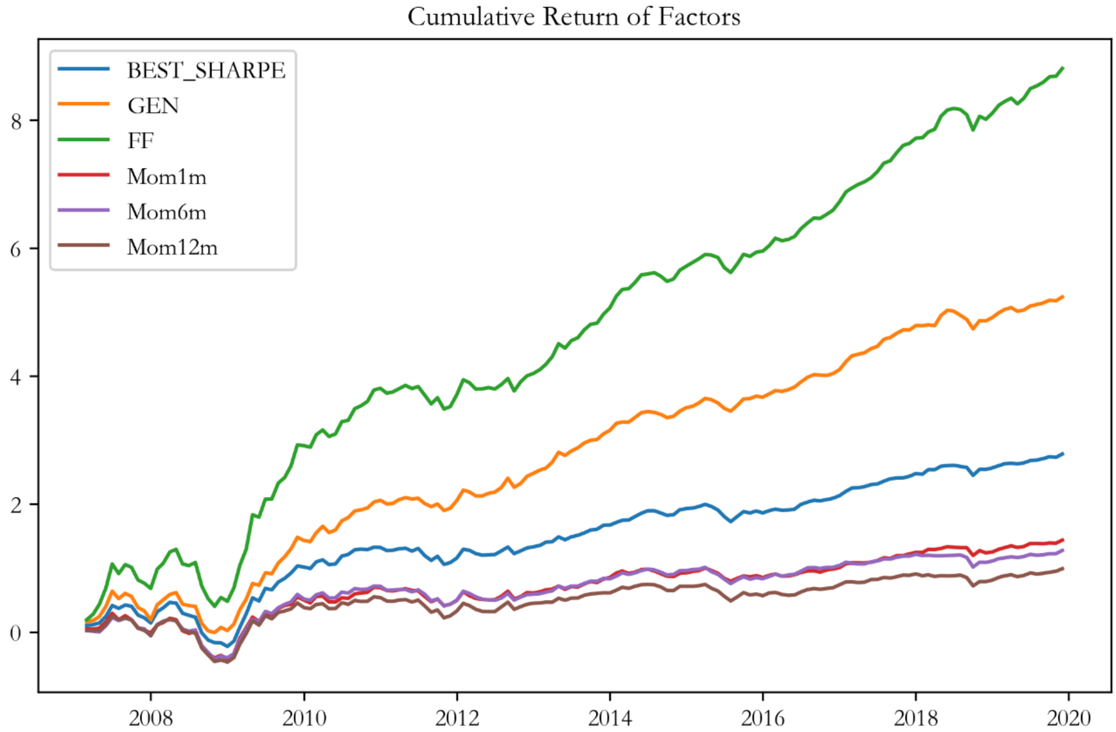

Figure 4-10 Comparison of Momentum-Like Factors

Next, the cumulative factor returns of quintile plots will show as follows. To clearly understand the direction of our factor ranking, we sort the factors from smallest to largest, with P1 representing the bottom 20% and P5 representing the top 20%. We rebalance the portfolios monthly and calculate the cumulative portfolio returns using equal weighting, observing the factor performance across different periods. The factor performance will show as follows.

Table 4-5 Factor Performance

Factor	Sharpe Ratio	Max DrawDown
TOP_10	0.96	48.83%
BEST_SHARPE	0.89	52.81%
BEST_SR	0.71	44.55%
GEN	1.54	54.52%
SME	1.67	49.25%
VALUE	1.23	52.54%
FF	1.88	64.86%
MOM1M	0.44	55.36%
MOM6M	0.37	52.60%
MOM12M	0.26	55.65%

We further observed the performance and drawdown resilience of each factor. We found that factors with a higher correlation to momentum tend to have poorer drawdown resilience, while factors with a lower correlation to momentum generally exhibited better drawdown performance, such as SME, TOP_10, and BEST_SR.

i. Factor Performance

TOP_10:

Based on Figure 4-7, we can observe that the cumulative return roughly follows the same trend as the factor ranking after 2014. Starting from 2014, the best-performing P1 consistently outperformed P5, with P1 reaching a peak of nearly 330% cumulative return. The cumulative Rank IC of -0.04 indicates that this factor ranking is negatively correlated with stock returns.

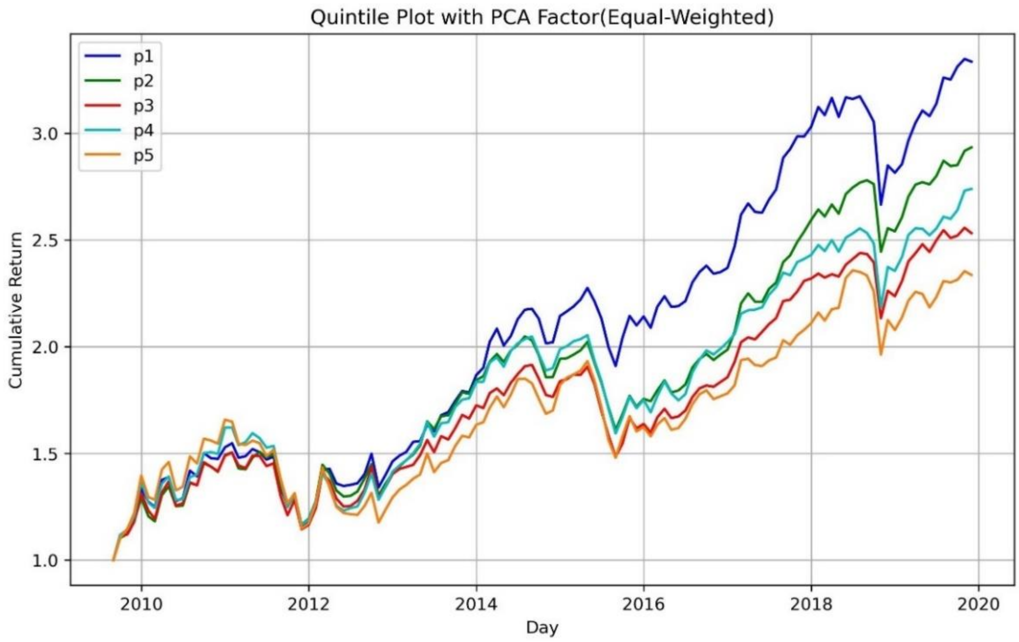

Figure 4-11 Cumulative Equal-Weighted Factor Returns by TOP_10

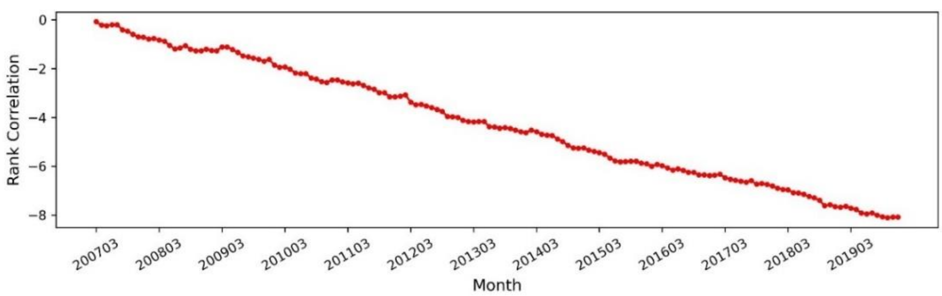

Figure 4-12 Cumulative Rank IC by TOP_10

BEST_SHARPE:

The ranking of this factor roughly correlates positively with returns, with the best-performing group, p5, showing the highest return of up to 250%, occurring in the last. Its cumulative Rank IC is 0.04, indicating some predictive power regarding stock returns. Additionally, there is a significant difference in performance between the p1 and p5 groups, suggesting that this factor may have a nice value-added effect.

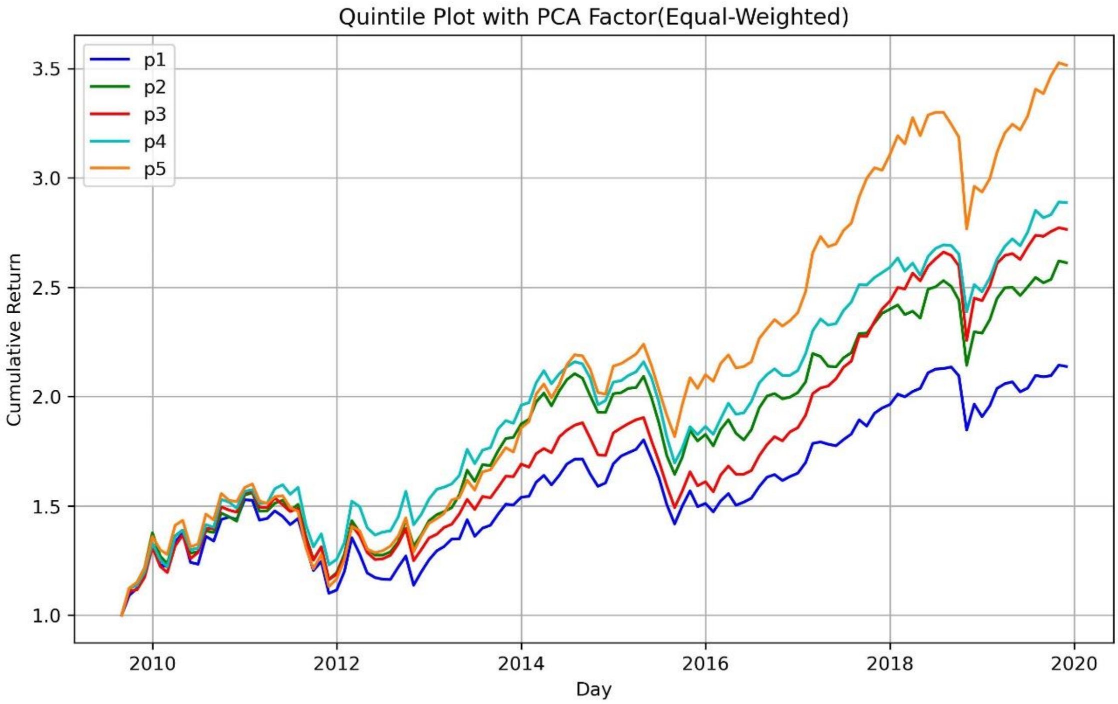

Figure 4-13 Cumulative Equal-Weighted Factor Returns by BEST_SHARPE

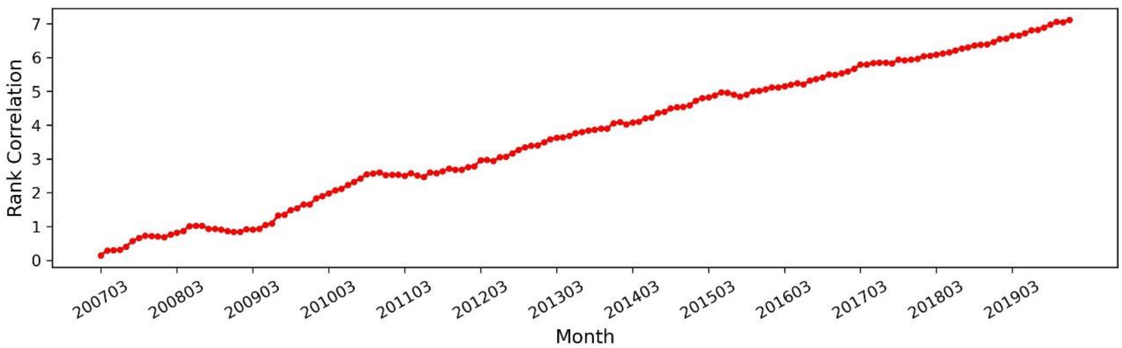

Figure 4-14 Cumulative Rank IC by BEST_SHARPE

BEST_SR:

Based on Figure 4-9, it can be observed that the performance of this factor in the p1 and p5 portfolios was quite similar before 2016 and both portfolios ranked in the top two in the sample period. The factor's cumulative Rank IC of 0.03 and the apparent volatility in its rank IC suggest that the effectiveness of this factor in value-added applications may be limited.

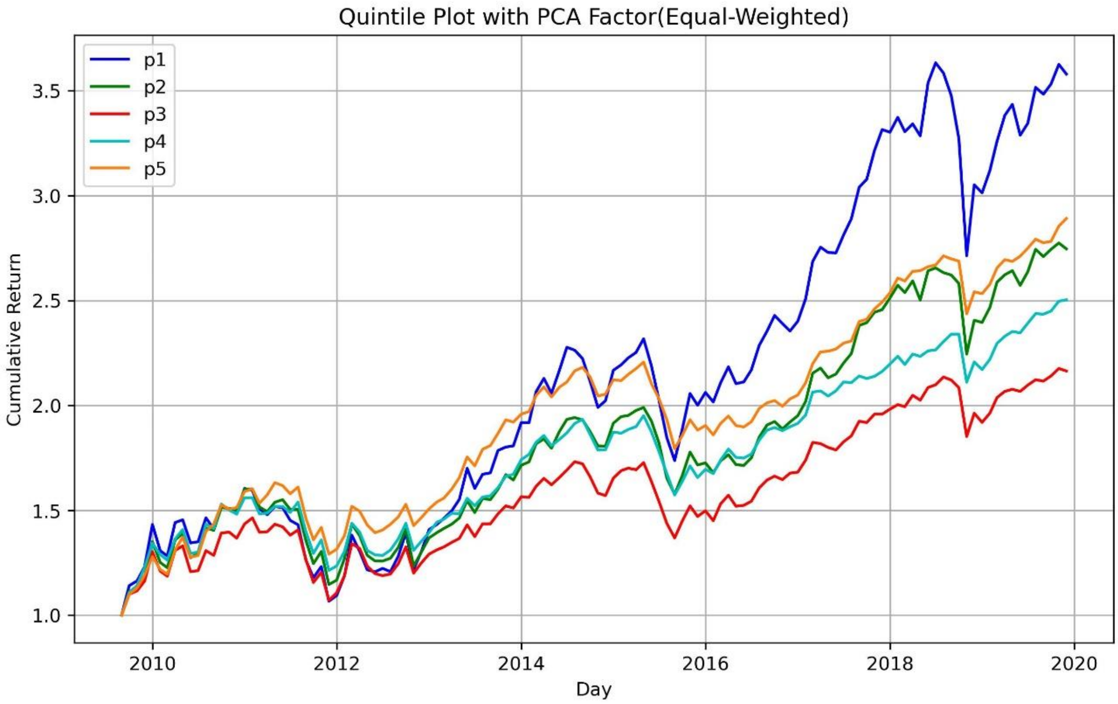

Figure 4-15 Cumulative Equal-Weighted Factor Returns by BEST_SR

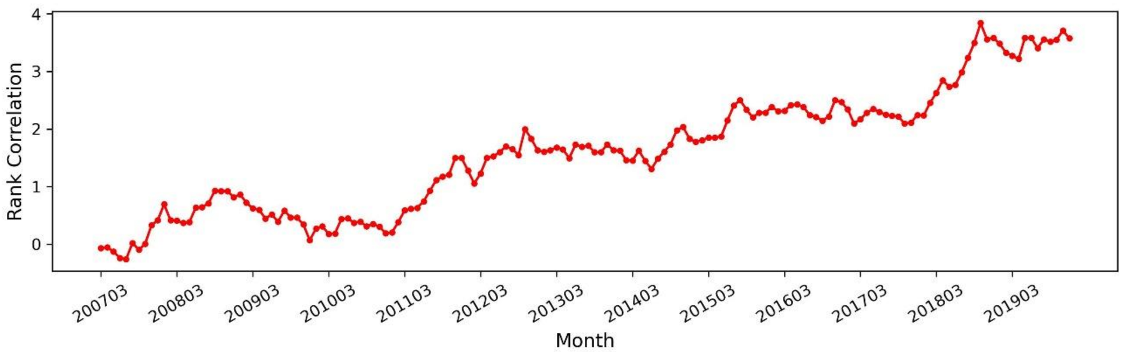

Figure 4-16 Cumulative Rank IC by BEST_SR

GEN:

The factor learned from GEN demonstrates excellent performance in the p5 portfolio, reaching a peak return of 500%, significantly outperforming other portfolios. Its cumulative rank IC is 0.058, traditionally considered effective when exceeding 0.05.

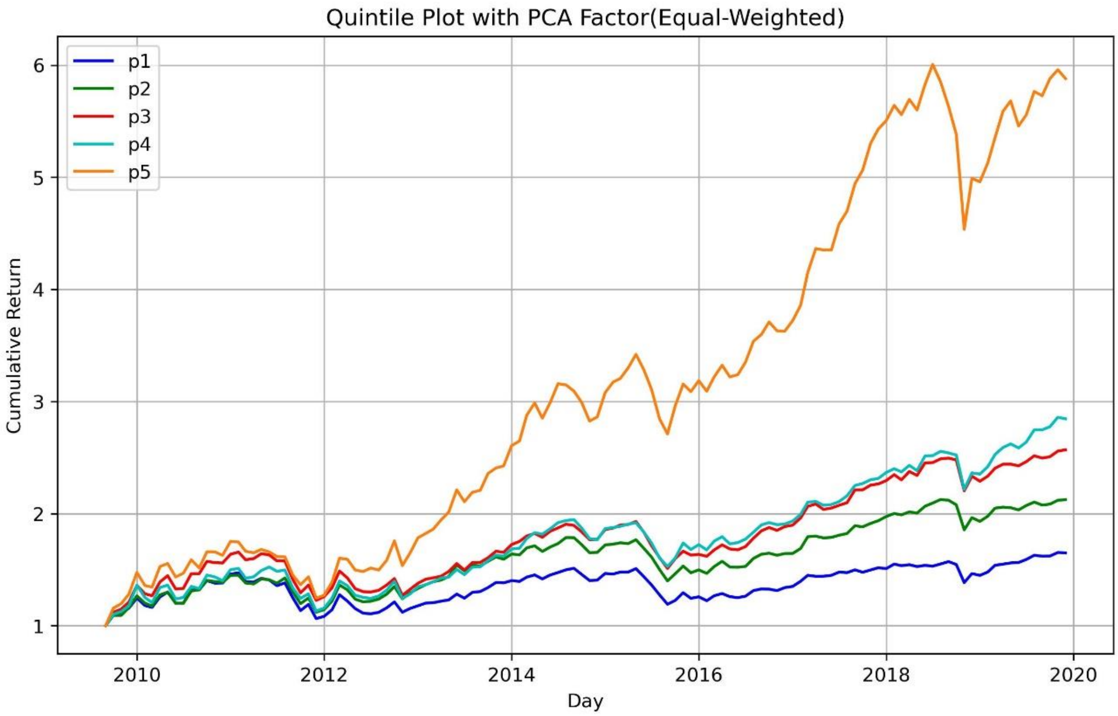

Figure 4-17 Cumulative Equal-Weighted Factor Returns by GEN

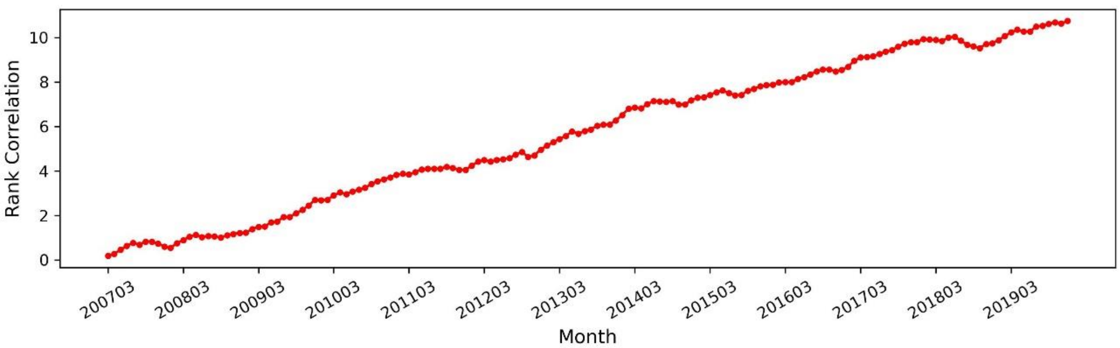

Figure 4-18 Cumulative Rank IC by GEN

SME:

The SME factor exhibits excellent performance, with the maximum drawdown of each portfolio being among the lowest of all deep factors. Its cumulative Rank IC is as high as 0.07, indicating that this is an outstanding factor. The performance of its value-added applications should be excellent.

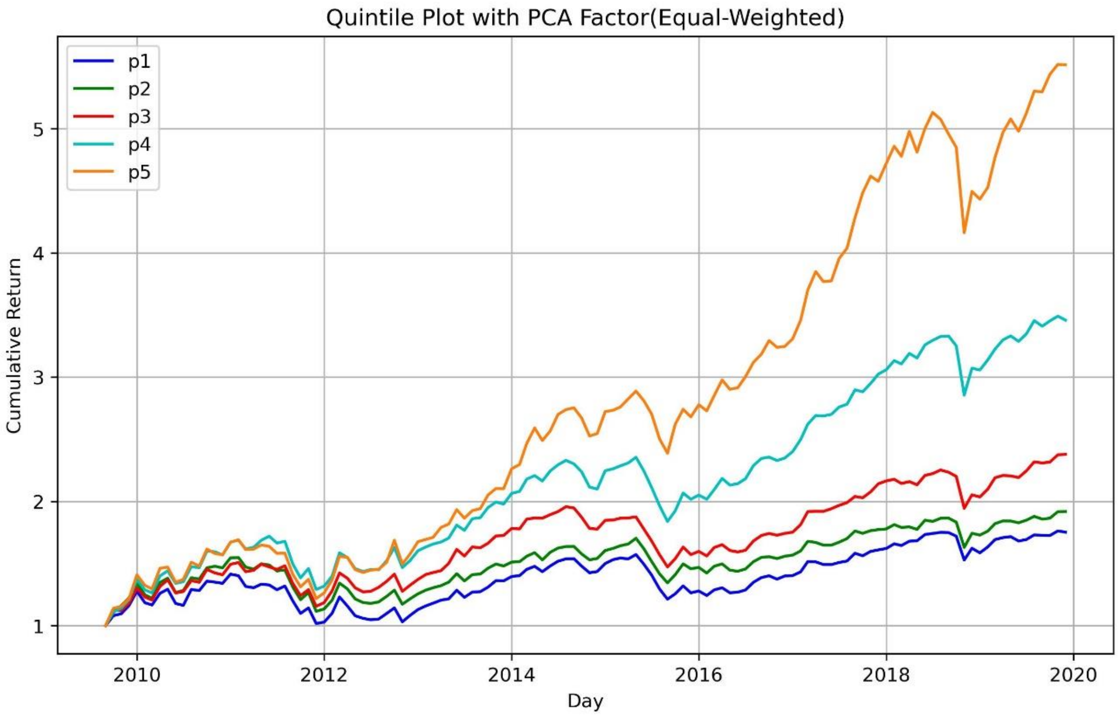

Figure 4-19 Cumulative Equal-Weighted Factor Returns by SME

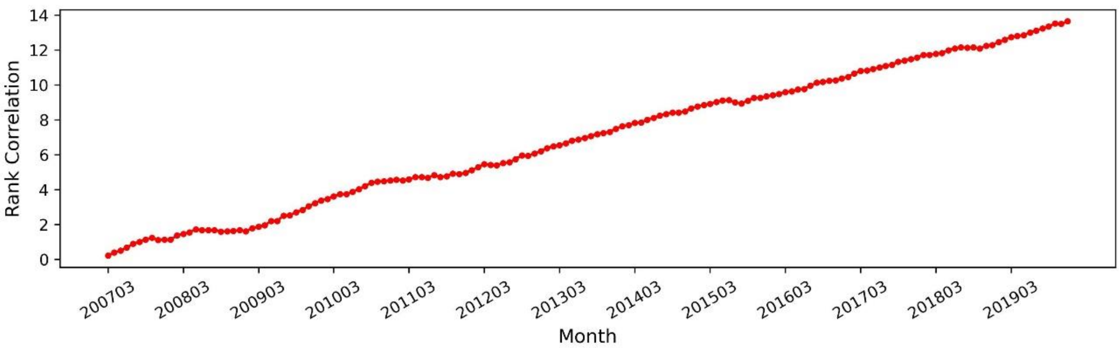

Figure 4-20 Cumulative Rank IC by SME

VALUE:

The cumulative return plot of the VALUE factor depicts relatively tangled performance across different portfolios, with a cumulative Rank IC of 0.01, the lowest among all deep factors. Traditionally, when a cumulative Rank IC value approaches 0, it may indicate that the factor has poor predictive power for forecasting future returns.

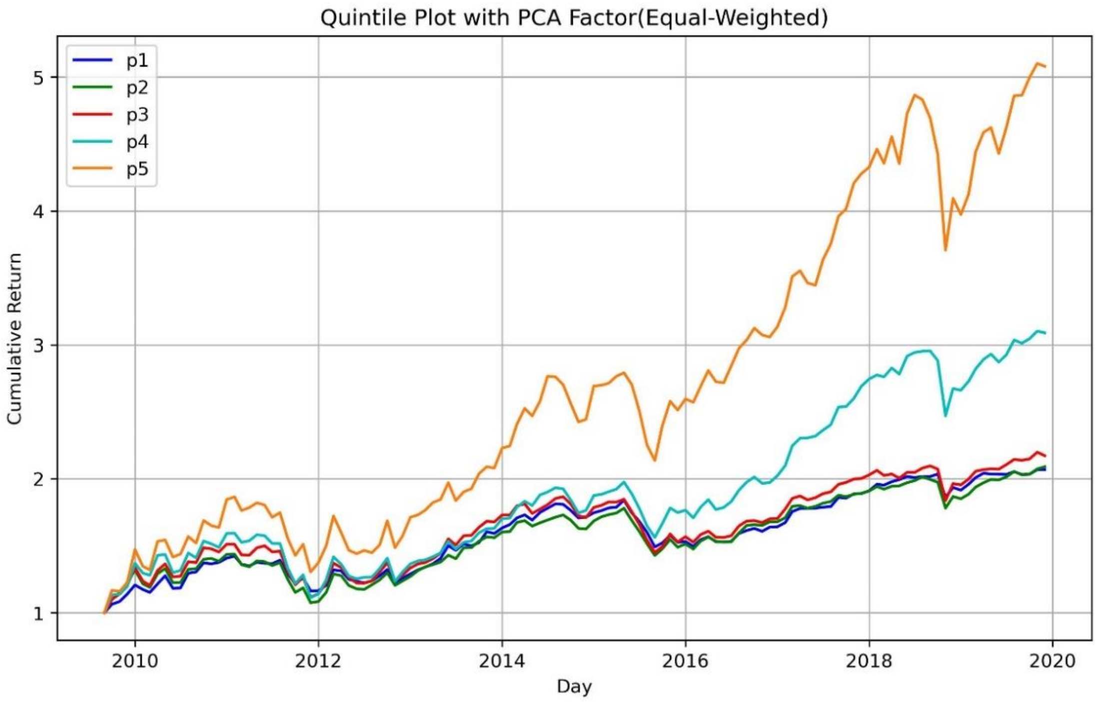

Figure 4-21 Cumulative Equal-Weighted Factor Returns by VALUE

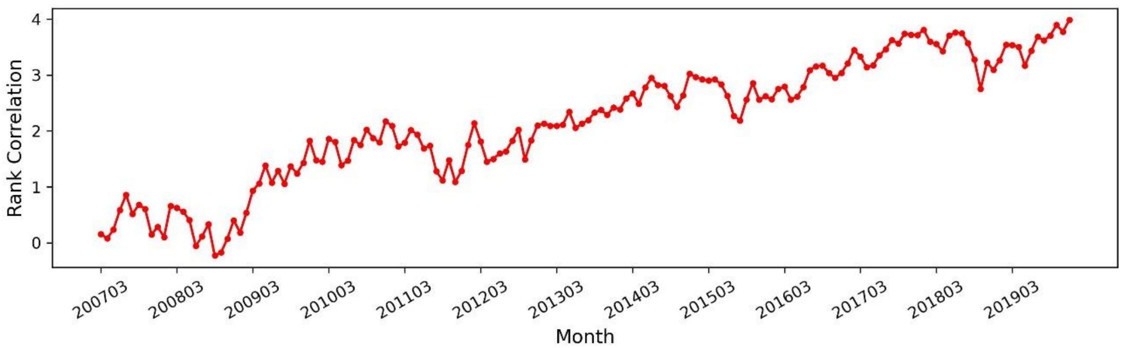

Figure 4-22 Cumulative Rank IC by VALUE

FF:

Throughout the entire sample period, p5 consistently outperformed other portfolios, while portfolios p1 through p4 maintained very similar performance. With a cumulative Rank IC of 0.046, this indicates that it is also a decent value-added factor.

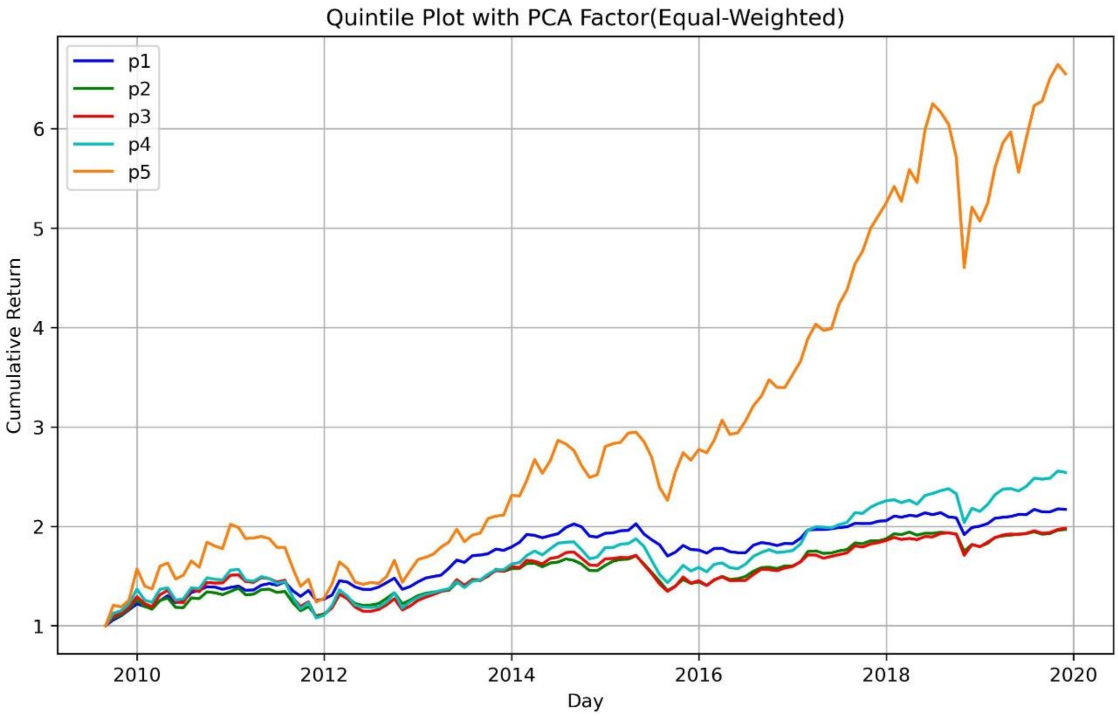

Figure 4-23 Cumulative Equal-Weighted Factor Returns by FF

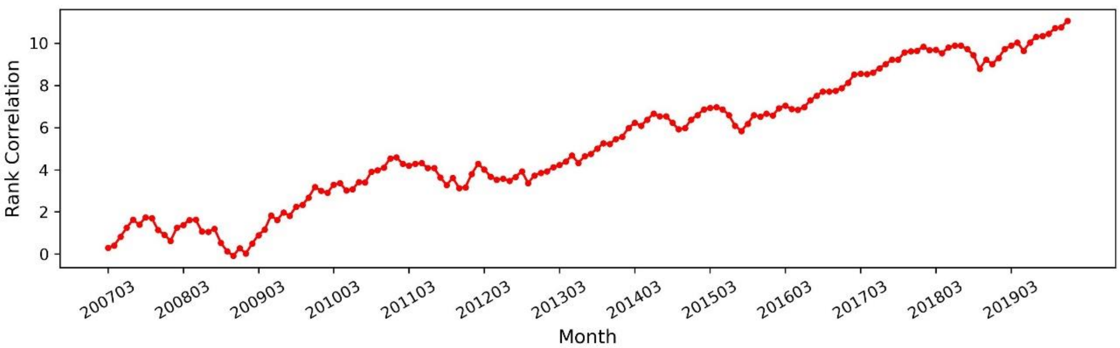

Figure 4-24 Cumulative Rank IC by FF

4.4 Risk Model Prediction Analysis

Before examining the predictive accuracy of the models, we need to evaluate the risk model's explanatory power. Figure 4-25 shows the explanatory power of different factor models across various periods. Although there is not much difference in their performance in the rolling 12-month chart, we can still use the average R-squared to assess the explanatory power of the models. Additionally, we have compiled all charts in Appendix D.

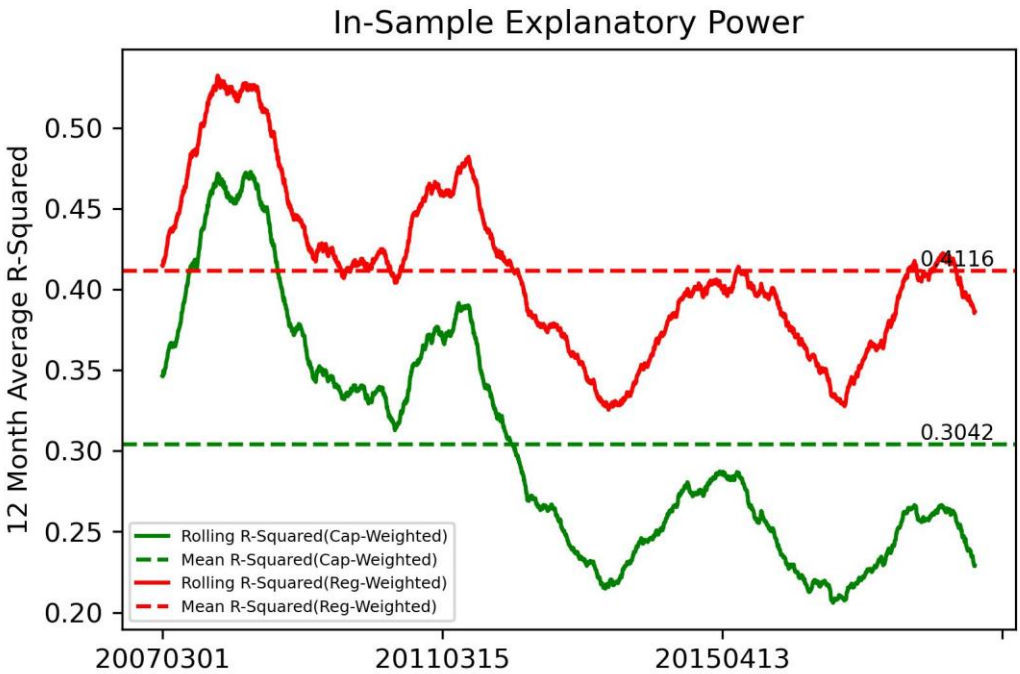

Figure 4-25 Explanatory Power of TOP_10

The cap-weighted R-squared ranges from 0.2 to 0.48, while the reg-weighted R-squared ranges from 0.32 to 0.55. The former weighting method ensures that companies with larger market capitalizations do not overly influence the explanatory power of the model. Adding more factors can easily increase the R-squared value. However, this may lead to overfitting. In this study, we mainly focus on the deep factors that we have constructed.

Finally, we apply the Bias test to examine the predictive ability of the risk models. These tests help determine whether the models can accurately forecast risk and if there are any consistent prediction biases. Figures 4-25 through 4-31 show the bias statistics over different periods and Table 4-5 provides a comprehensive summary of the results from bias tests for all factors.

In the Bias test, a B-Value of 1 indicates accurate prediction, a B-Value greater than 1 suggests undervaluation risk and a B-Value less than 1 indicates overvaluation risk. In summary, we observe that most models exhibit a slight tendency towards undervaluation risk.

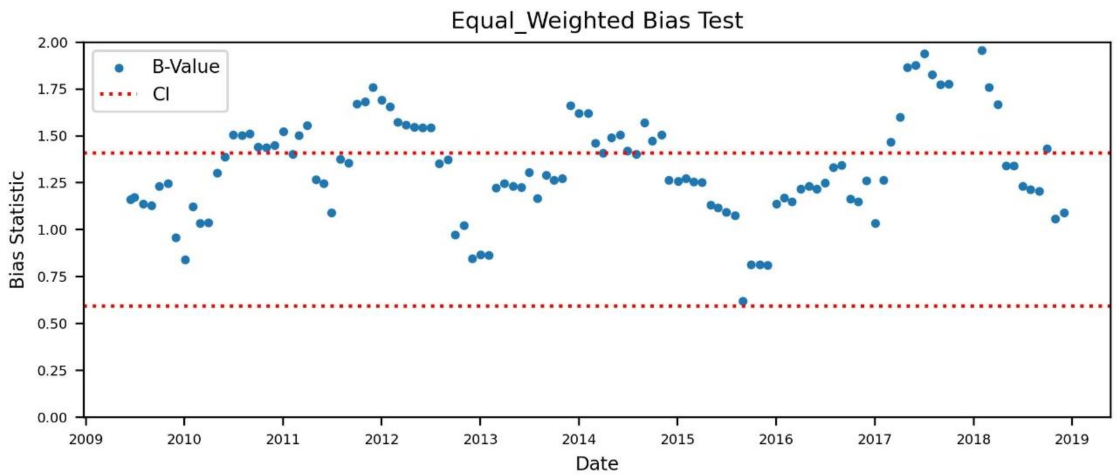

Figure 4-26 Bias Test of TOP_10

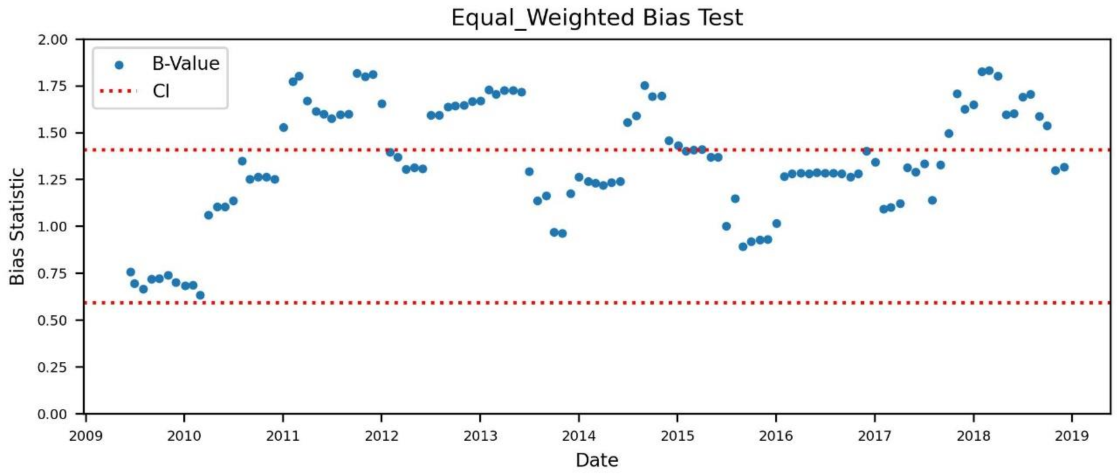

Figure 4-27 Bias Tests of BEST_SHARPE

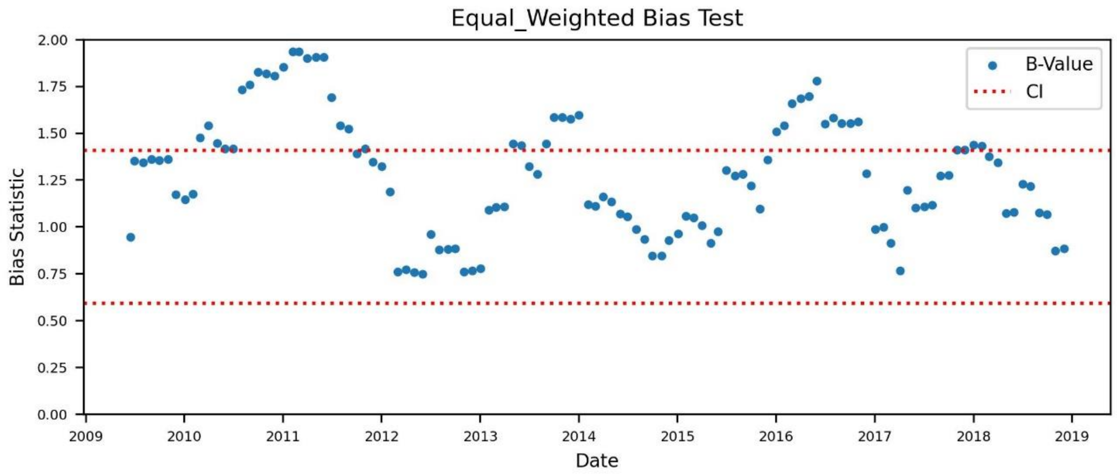

Figure 4-28 Bias Test of BEST_SR

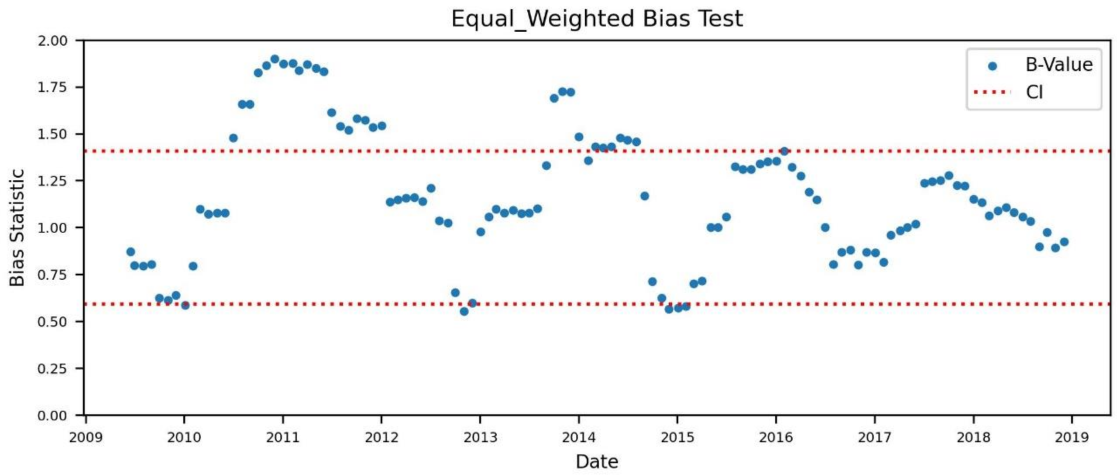

Figure 4-29 Bias Test of GEN

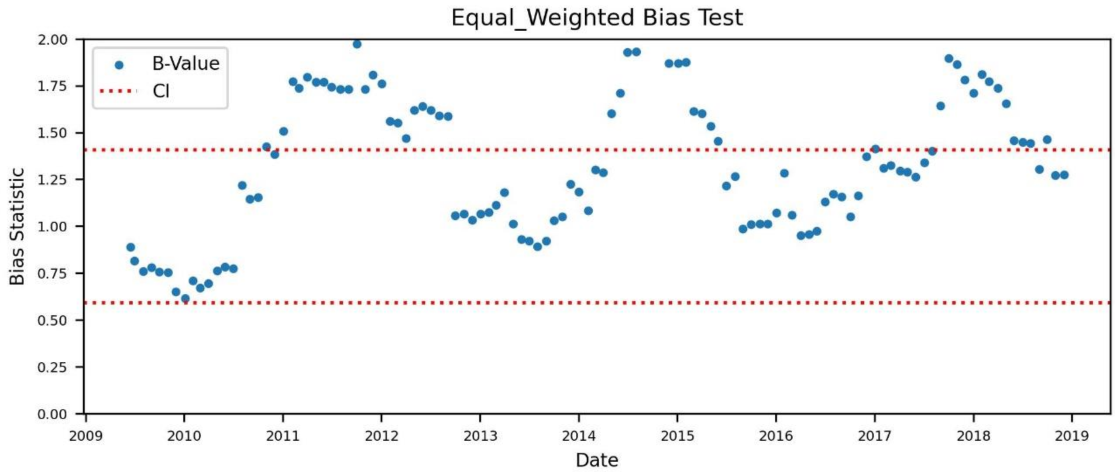

Figure 4-30 Bias Test of SME

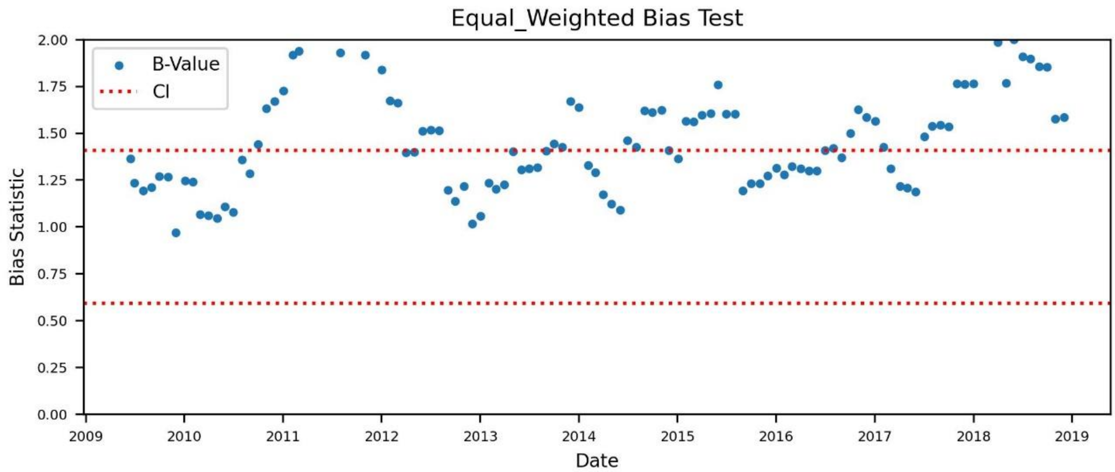

Figure 4-31 Bias Test of VALUE

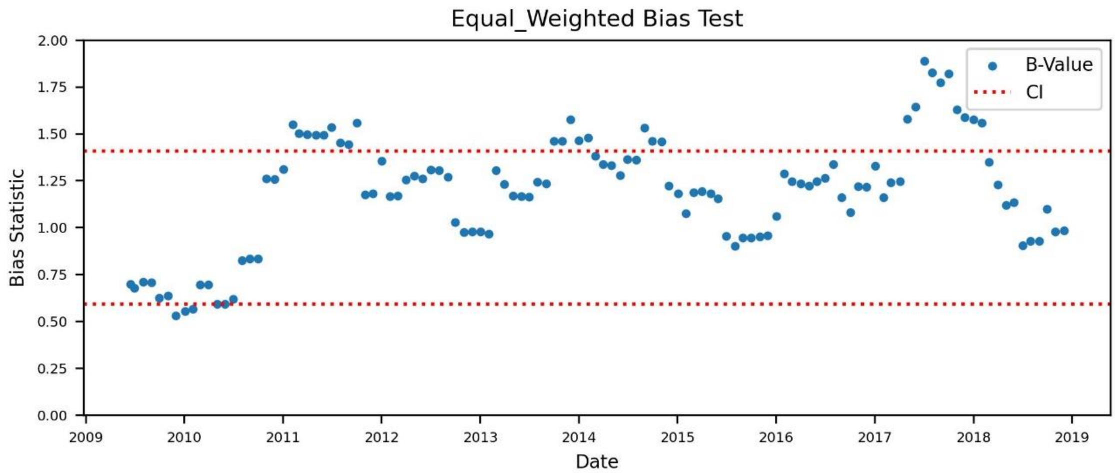

Figure 4-32 Bias Test of FF

Table 4-5 displays the proportion of B-Values falling within the confidence level. According to empirical rules, a prediction accuracy of 85% or higher is well-suited for application in stock markets, assisting investors in risk planning.

Table 4-6 Bias Test Results

Factor	Robust	Without Robust
TOP_10	67.83%	64.35%
BEST_SHARPE	66.09%	58.26%
BEST_SR	74.78%	67.83%
GEN	80.00%	72.17%
SME	82.61%	67.83%
VALUE	66.09%	58.26%
FF	93.04%	80.87%

Of all factors, FF shows the strongest risk prediction capability, with SME and GEN also displaying relatively strong abilities in this regard. However, other factors exhibit weaker performance. Our findings suggest that factors with higher cumulative Rank IC generally possess superior risk prediction capabilities.

4.5 Investment Portfolio

We constructed tracking portfolios for the Taiwan 50 Index based on the index ground rule. The sample period extended from August 2009 to March 2020.

i. Taiwan 50 Index Portfolio

We applied the full replication method to track the Taiwan 50 Index. Using 50 stocks to track all 50 stocks, in cases where individual stocks have no trading data, experience cash increases/decreases, or are delisted, etc., additional companies from around the 50^{th} to 55^{th} position based on market capitalization are included to maintain the total number of constituent stocks. Rebalancing takes place quarterly. The tracking performance is presented in Table 4-5. The tracking error over the entire sample period is 0.16%. According to the table, except for 2010 where there is a higher tracking error of 0.37%, the tracking error remains consistently low throughout the sample period. The total tracking difference is -0.31%, which shows that the tracking portfolio has slightly underperformed the benchmark index.

Table 4-7 Taiwan 50 Index Tracking Performance

Year	Tracking Error	Tracking Difference
2009	0.11%	-0.01%
2010	0.37%	-0.32%
2011	0.10%	-0.10%
2012	0.08%	0.11%
2013	0.05%	-0.06%
2014	0.15%	0.07%
2015	0.09%	0.01%
2016	0.11%	0.07%
2017	0.09%	0.05%
2018	0.24%	0.26%
2019	0.12%	-0.38%

To better observe the alpha generation ability of the factors, we tried to push the tracking error close to 4% and calculated the annualized active risk and active return. The results are as follows:

Table 4-8 Enhanced Performance with k=1(Annualized)

Factor	Active Risk	Active Return	IR
TOP_10	0.55%	-0.09%	-0.16
BEST_SHARPE	0.53%	-0.06%	-0.11
BEST_SR	0.71%	0.04%	0.05
GEN	0.59%	0.05%	0.09
SME	0.53%	-0.22%	-0.42
VALUE	0.71%	-0.16%	-0.23
FF	0.85%	0.08%	0.09
NONE	0.17%	-0.11%	-0.66

Table 4-9 Enhanced Performance with Active Risk of 4%(Annualized)

Factor	Active Risk	Active Return	IR
TOP_10	3.41%	2.44%	0.72
BEST_SHARPE	3.57%	3.17%	0.89
BEST_SR	3.85%	3.11%	0.81
GEN	3.71%	3.67%	0.99
SME	4.00%	3.77%	0.94
VALUE	3.75%	3.49%	0.93
FF	3.08%	1.99%	0.65

For each adjustment, increment k by 0.1 until active risk breaks through the previous value of 4%. The transition cost is set at 0.3%.

From Table 4-7, we can observe that when we set the total adjusted weight at 5%, TOP_10, BEST_SHARPE, SME, and FF exhibit better performance, with higher cumulative Rank IC. We then attempted to push the active risk close to 4%, and each portfolio still showed commendable value-addition performance.

Therefore, we further examine the alpha generation effect through a cumulative return comparison of the benchmark and the enhanced portfolio in the following Figure.

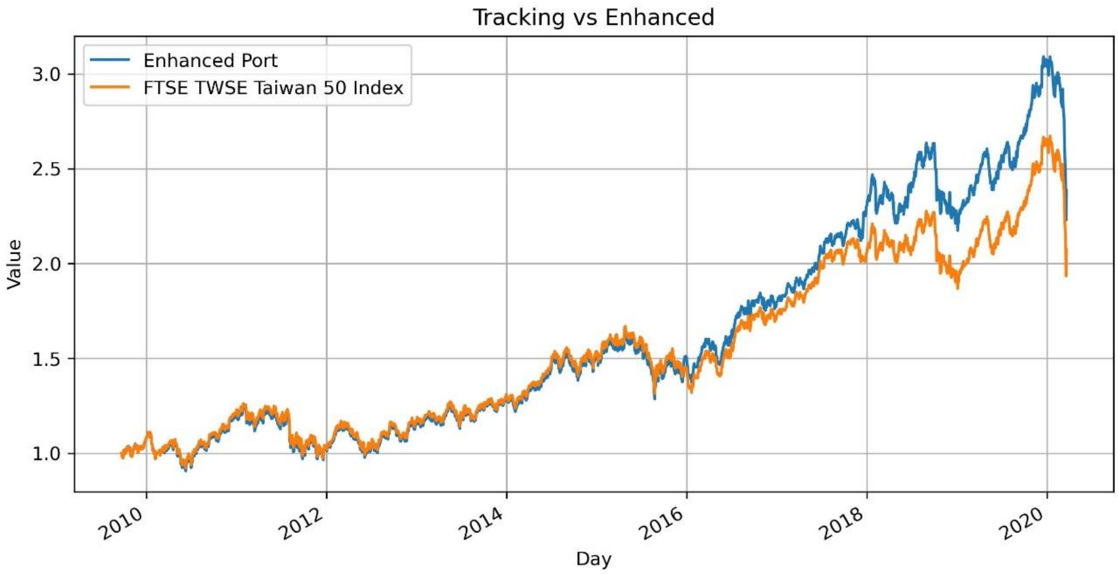

Figure 4-33 TOP_10 Portfolio Comparison

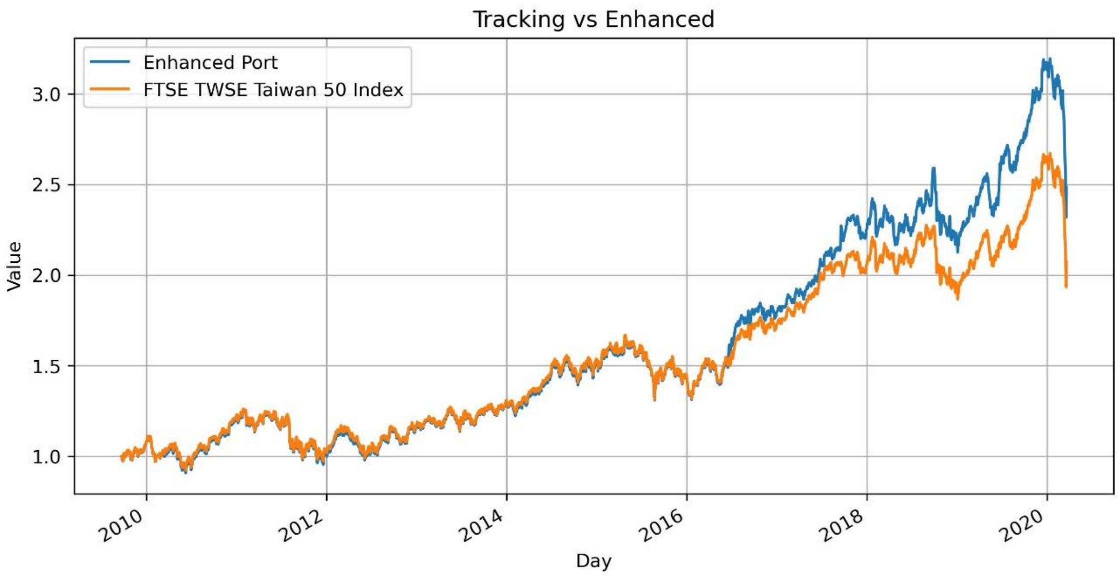

Figure 4-34 BEST_SHARPE Portfolio Comparison

TOP_10 and BEST_SHARPE exhibit similar alpha generation effects, with both factors producing more value-enhancing benefits starting around 2016.

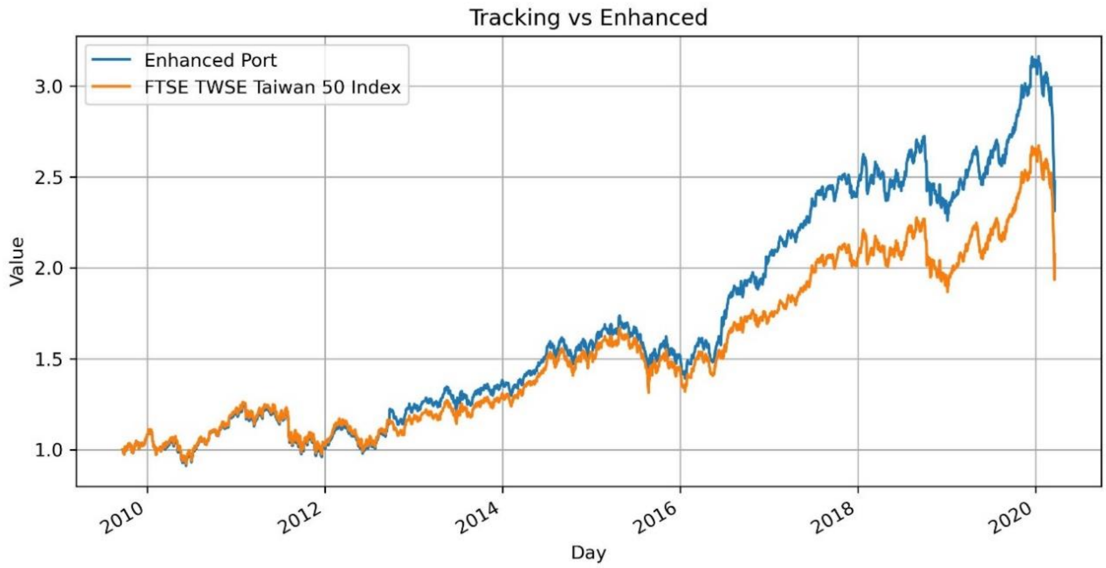

Figure 4-35 BEST_SR Portfolio Comparison

BEST_SR started demonstrating stable value-enhancing effects in 2013 and showed excellent effects from 2016 to 2020. We found that GEN showed the best value-enhancing performance among all factors. According to our value enhancement approach, the performance of GEM in the quintile plot reveals a significant difference in performance between the highest and lowest factor exposures when compared to other factors.

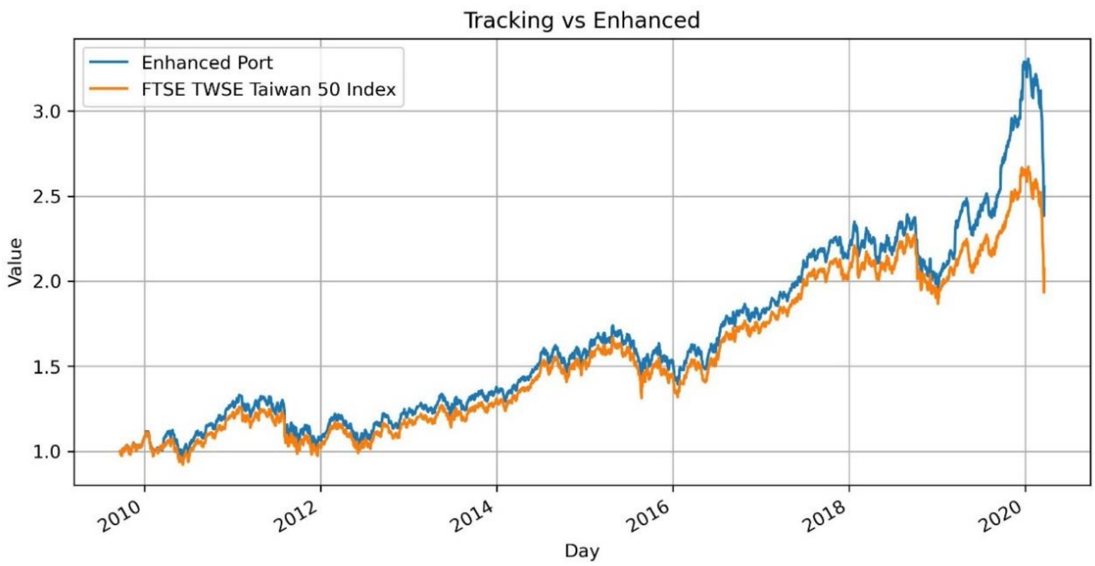

Figure 4-36 GEN Portfolio Comparison

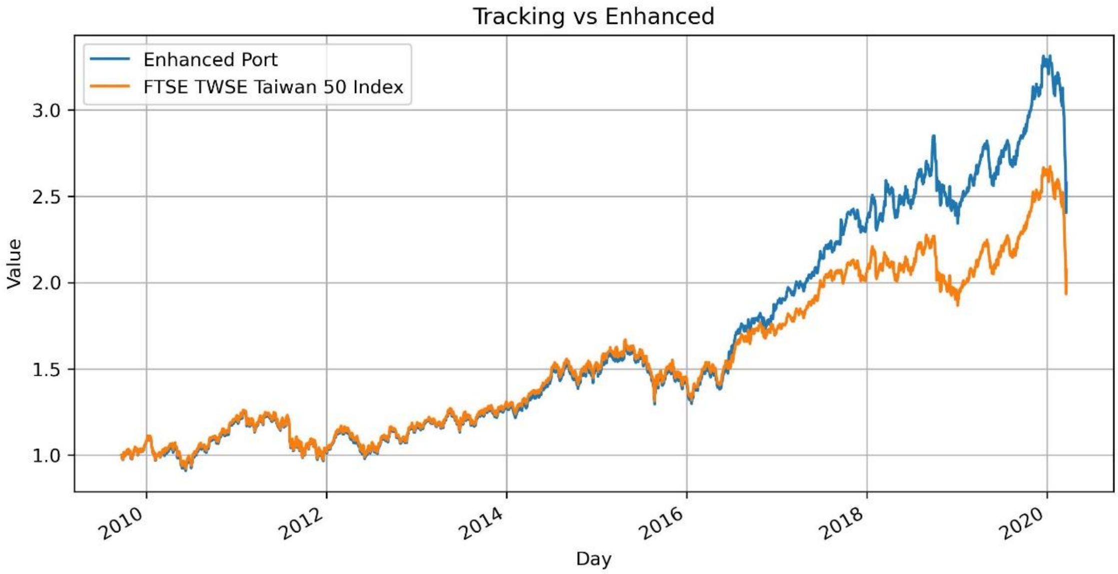

Figure 4-37 SME Portfolio Comparison

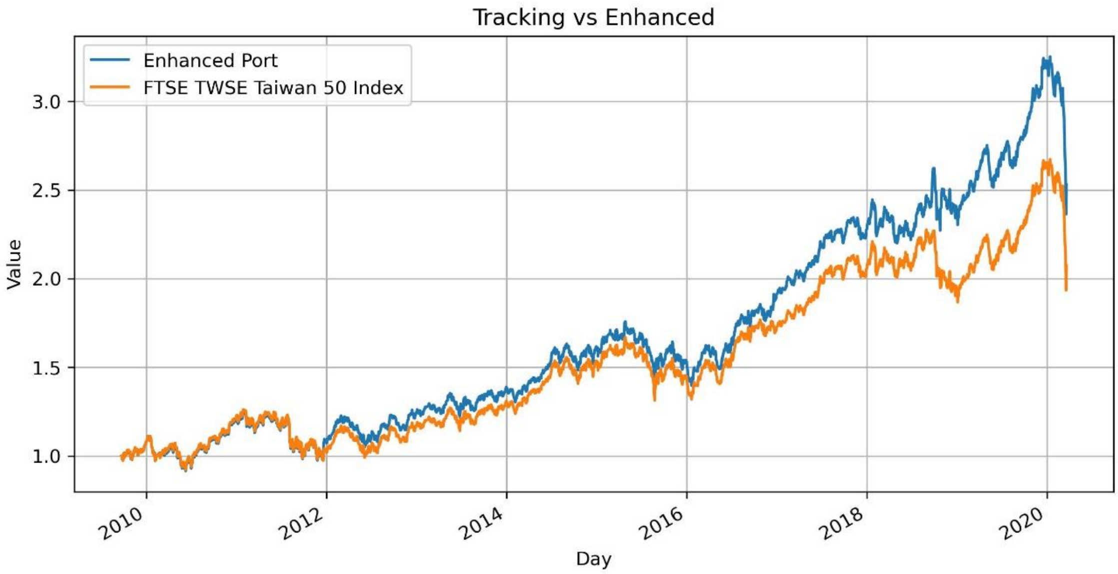

Figure 4-38 VALUE Portfolio Comparison

SME and VALUE have information ratios of 0.94 and 0.93, respectively. According to the quintile plots for the two factors, the distribution of cumulative returns between the highest and lowest factor exposures is similar. Although SME surpasses VALUE significantly in cumulative Rank IC, since the value enhancement only considers the stocks with the highest and lowest factor exposures, the value-added effects of both factors appear very similar over the sample period.

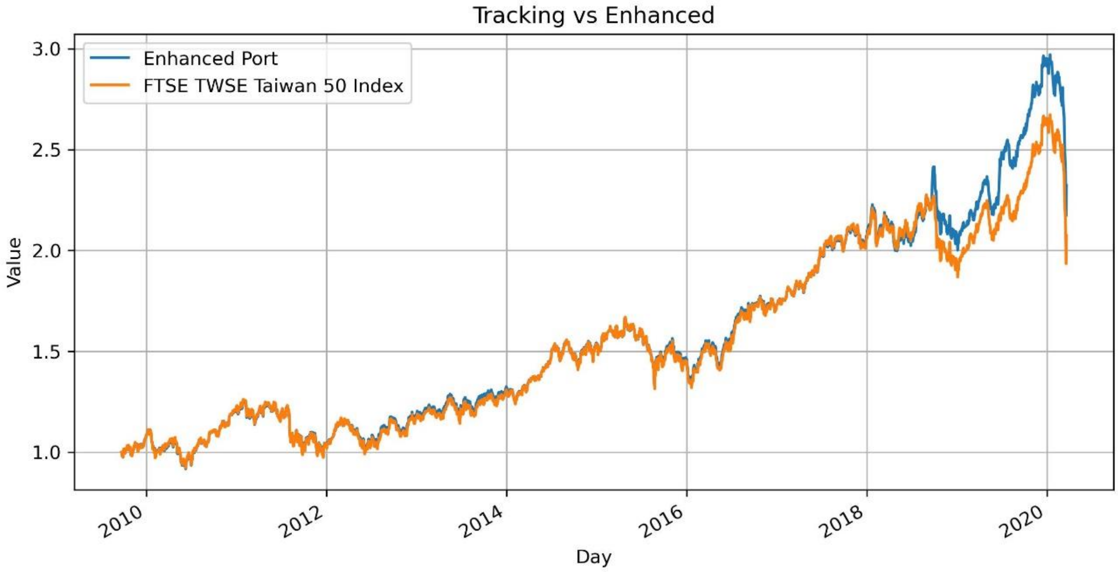

Figure 4-39 FF Portfolio Comparison

FF's value-added performance is relatively inferior to other factors, and its best-performing period also corresponds to the lowest returns. We can infer that the deep factor constructed by FF is relatively less applicable to the Taiwan stock market compared to the other target portfolios.

By employing different Target Portfolios learned from deep factors, we can see the factors do a good job in terms of annualized Alpha and IR as the active risk approaches 4%. For the value-added effect, we discover different trends in terms of the enhancement over time and that's mostly owed to the fact that the Alpha computed as the difference of the end and start values of the portfolio. This said we saw that GEN and VALUE have always shown stable and superior enhancement than all other portfolios.

In addition to the value-enhancing benefits of the portfolio, we further explore the application of factors in the Long-Short strategy. To diversify the risk, we implement a 130/30 strategy, wherein we allocate equal investments to the top 30 and bottom 15 stocks based on factor exposure. This portfolio is rebalanced monthly. The performance is as follows:

Table 4-10 Long-Short Strategy Performance(Annualized)

Factor	Return	Risk	Sharpe Ratio	MDD
TOP_10	23.13%	20.77%	1.11	58.51%
BEST_SHARPE	24.94%	25.47%	0.98	73.09%
BEST_SR	6.25%	15.06%	0.42	52.17%
GEN	67.18%	34.20%	1.96	56.25%
SME	59.38%	32.16%	1.85	61.78%
VALUE	77.10%	56.07%	1.38	57.96%
FF	120.03%	62.42%	1.92	99.57%

The transition cost is set at 0.3%. MDD represents max drawdown.

GEN, SME, and FF have a high Sharpe ratio, but FF has a higher max drawdown, suggesting that GEN and SME may be the two best-performing factors in the Long-Short strategy.

5. Conclusions and Suggestions

5.1 Conclusions

This study focuses on utilizing MLP deep learning models to learn the process of factor ranking in funds and applies the learned results to factor strategies, risk analysis, and portfolio strategies. The key advantages of MLP are (1) End-to-end learning capability, avoiding manual feature engineering biases. (2) Flexible model scalability to handle tasks of varying complexities. (3) Strong generalization ability for excellent performance on unseen data.

Building upon the original framework proposed by Feng et al. (2023), this study further optimizes by incorporating style analysis to construct target portfolios. Different from traditional indicator-based sorting methods, this enhancement allows the overall framework to be applied to learning the factor ranking process of various funds, such as Normal funds and Index funds.

Additionally, parameter adjustments are made based on the characteristics of the Taiwan stock market. This study addresses each potential pitfall highlighted in Cao (2023) for the application of machine learning and deep learning in financial markets, including (1) Possible Overfitting, by incorporating random dropout and parameter optimization. The train and test losses showed no signs of overfitting. (2) Implementation Gap, by considering trading costs in the analysis of trading strategies, and controlling for stock liquidity (selecting large-cap listed companies). (3) Explainability and Performance Attribution, by checking for multicollinearity issues among firm characteristics and removing 5 features with high VIF, running Fama-MacBeth regressions to identify important features from the deep learning model outputs. In the empirical results, we find that the factor performance learned from Taiwan stock market funds significantly outperforms the best-performing model in Feng et al. (2023), demonstrating that the learning objectives presented in this study are more suitable for the Taiwan stock market, which is consistent with our expectations.

Most models demonstrate commendable asset pricing ability, with Trad R-squared values exceeding 70% during the out-of-sample period, except for FF. When controlling for the Fama-French three factors to measure asset pricing ability, VALUE exhibits the best performance.

In single-factor analysis, a factor with positive returns most of the time and low volatility is considered an excellent return factor. A factor that can significantly describe a certain systematic risk is regarded as an excellent risk factor. SME and GEN are excellent return factors, exhibiting cumulative Rank IC values above 0.05. The FF, which performed best in Feng et al. (2023), is an outstanding risk factor. In the application of portfolio enhancement, each factor demonstrates promising performance. Particularly, SME, GEN, and VALUE exhibit superior performance. In the Long-Short strategy, SME, GEN, and FF demonstrate Sharpe Ratios surpassing 1.85, suggesting their commendable risk-adjusted returns. However, it's worth noting that FF demonstrates higher overall annualized risk and a Max DrawDown approaching 100%, contravening acceptable investment standards.

Due to the deep learning framework, which utilizes learned factors alongside the Fama French 3 factors to assess model explanatory power (asset pricing ability), we observe that the learned factors capture more momentum characteristics. Furthermore, upon closer examination, we find that combining momentum features with certain other factor categories effectively mitigates the issue of momentum crashes. We found that if a factor's components are highly correlated with the pure momentum factor, then during a momentum crash period, this factor will not be able to effectively mitigate the decline. If a factor's components have a lower correlation with the pure momentum factor, then during a momentum crash period, this factor can further mitigate losses. By combining various types of features to construct a factor, even if most of the factor's components are momentum-type features, it can significantly reduce the correlation between that factor and the pure momentum factor.

In addition to the aforementioned sample period, this study also tested the model's performance using different sample periods to observe its performance between January 2020 and December 2022. During this period, events like the COVID-19 pandemic, the Federal Reserve's aggressive interest rate hikes, and the Russo-Ukrainian War occurred. These events had significant impacts on the financial markets. Therefore, evaluating the model's performance during these turbulent times is crucial for assessing its resilience and predictive capabilities. The overall learning performance, as measured by the R-squared, experienced an average decline of 10% during the out-of-sample period. When using the Taiwan stock market funds as the target portfolio, the average Trad R-squared was 65%, while for FF, it was 55%. This suggests that periods containing numerous financial storms may not be as suitable for use as out-of-sample periods. Regarding asset pricing ability, the model performance is still praiseworthy.

In conclusion, the main contributions of this research are three: (1) This deep learning architecture can be flexibly applied to analysis applications across various funds. For different markets, the required input structure can also be constructed through style analysis. (2) We find that using general and small-mid cap funds in the Taiwan stock market as learning targets can yield better factors. (3) Combining momentum factors with different categories of factors can effectively mitigate the momentum crash issue.

5.2 Suggestions

First, this paper focuses on the ranking process of learning factors. After analyzing the results obtained from the deep learning model, we can further apply these factors in various ways. The following research can involve incorporating different deep learning models based on different application scenarios, such as integrating the LSTM model for entry and exit signal determination, which is good at capturing long-term dependencies in time series data. Utilizing a combination of various models can enhance the optimization of trading strategies. Due to space constraints, this research does not present the related empirical results. However, we have conducted corresponding research, utilizing chip data in conjunction with an LSTM model to determine entry and exit timing. Empirical findings revealed that this can further enhance the strategy's Sharpe Ratio and reduce the Maximum Drawdown, but since the determination of entry and exit timing in strategy implementation often requires substantial position adjustments, it may lead to a higher turnover rate. For example, by using chip data to determine upward trends as buy/sell signal indicators, we found that the net buying/selling positions of the three major institutional investors were an important feature for trend identification. After incorporating an LSTM model for entry and exit signal determination on the BEST_SHARPE factor, the Sharpe ratio of the long-short strategy improved from 0.98 to

1.05, and the maximum drawdown decreased from 73.9% to 67.43%.

Second, based on the data obtained in this study, financial data appeared at a semi-annual frequency before January 2007 and switched to a quarterly frequency after January 2007. Due to the difference in data frequency, constructing characteristics requires addressing numerous details. Therefore, the sample period of this study is set from January 2007 to December 2019, with a sample split ratio of 0.76:0.24 for in-sample and out-sample data. Increasing the dataset size may lead to better model generalization effects.

Third, this study only focuses on fundamental and technical features. Incorporating additional variables could potentially improve the model's robustness and adaptability in handling extreme events, leading to a more comprehensive assessment. For instance, (1) Market Sentiment Indicators, which can reflect investor panic, such as the investor sentiment index or news sentiment analysis; and (2) Macroeconomic Indicators, which can capture events that may lead to significant market fluctuations, such as GDP growth rate and inflation rate. With the advancements in computing power, processing higher-dimensional data is no longer as inefficient as before.

Finally, as time progresses, markets may become more efficient, and the alpha generated by specific factors could gradually transition to beta. Continuously optimizing parameters to adapt the model to changing market conditions would be a prudent choice. Additionally, incorporating machine learning techniques, such as reinforcement learning, could enable the model to dynamically adjust to evolving market dynamics. Conducting periodic reviews of the model's performance and recalibrating the feature set based on recent data will maintain its relevance and accuracy. Furthermore, ensuring transparency in the modeling process and maintaining a robust risk management framework is essential to mitigate potential model risks and achieve reliable outcomes.

References

1. Abe, M., & Nakayama, H. (2018). Deep learning for forecasting stock returns in the cross-section. Advances in Knowledge Discovery and Data Mining: 22nd Pacific-Asia Conference, PAKDD 2018, Melbourne, VIC, Australia, June 3-6, 2018, Proceedings, Part I 22,
2. Ahmed, S., Bu, Z., & Tsvetanov, D. (2019). Best of the best: A comparison of factor models. *Journal of Financial and Quantitative Analysis*, *54*(4), 1713-1758.
3. Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *The Journal of Finance*, *68*(3), 929-985.
4. Biner, B., Smith, R. C., & Ward, P. (2009). The Barra Europe Equity Model (EUE3). *MSCI Research Notes*.
5. Cao, L. (2023). *Handbook of Artificial Intelligence and Big Data Applications in Investments*. CFA Institute Research Foundation.
6. Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.
7. Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, *51*(5), 1681-1713.
8. Chen, L., Pelger, M., & Zhu, J. (2024). Deep learning in asset pricing. *Management Science*, *70*(2), 714-750.
9. Chen, W.-W. (2019). *Applying the Stock Chart Pattern Recognition with Deep Learning to Construct the Optimal Investment Strategy in Taiwan*
10. Cong, L. W., Feng, G., He, J., & He, X. (2023). *Asset pricing with panel tree under global split criteria*.
11. Connor, G. (1995). The three types of factor models: A comparison of their explanatory power. *Financial Analysts Journal*, *51*(3), 42-46.
12. Cooper, M. J., Gutierrez Jr, R. C., & Hameed, A. (2004). Market states and momentum. *The Journal of Finance*, *59*(3), 1345-1365.
13. Cox, J. C., & Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, *3*(1-2), 145-166.
14. Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, *122*(2), 221-247.
15. Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The Journal of Investing*, *11*(3), 7-22.
16. Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The Journal of Finance*, *50*(1), 131-155.
17. Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of*

Financial Economics, *116*(1), 1-22.

18. Feng, G., He, J., Polson, N., & Xu, J. (2023). Deep Learning in Characteristics-Sorted Factor Models. *Journal of Financial and Quantitative Analysis, Forthcoming*.
19. Foster, G., Olsen, C., & Shevlin, T. (1984). Earnings releases, anomalies, and the behavior of security returns. *Accounting Review*, 574-603.
20. Green, J., Hand, J. R., & Zhang, X. F. (2013). The supraview of return predictive signals. *Review of Accounting Studies*, *18*, 692-730.
21. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223-2273.
22. Gu, S., Kelly, B., & Xiu, D. (2021). Autoencoder asset pricing models. *Journal of Econometrics*, *222*(1), 429-450.
23. Gupta, F., Prajogi, R., & Stubbs, E. (1999). The information ratio and performance. *Journal of Portfolio Management*, *26*(1), 33.
24. Hanauer, M. X., & Kalsbach, T. (2023). Machine learning and the cross-section of emerging market stock returns. *Emerging Markets Review*, *55*, 101022.
25. Hill, J. M., & Naviwala, H. (1999). Synthetic and enhanced index strategies using futures on US indexes. *Journal of Portfolio Management*, 61.
26. Hou, K., Mo, H., Xue, C., & Zhang, L. (2021). An augmented q-factor model with expected growth. *Review of Finance*, *25*(1), 1-41.
27. Hou, K., & Robinson, D. T. (2006). Industry concentration and average stock returns. *The Journal of Finance*, *61*(4), 1927-1956.
28. Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, *28*(3), 650-705.
29. Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies*, *33*(5), 2019-2133.
30. 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.
31. Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, *134*(3), 501-524.
32. Kozak, S., Nagel, S., & Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, *135*(2), 271-292.
33. Leippold, M., Wang, Q., & Zhou, W. (2022). Machine learning in the Chinese stock market. *Journal of Financial Economics*, *145*(2), 64-82.
34. Leung, E., Lohre, H., Mischlich, D., Shea, Y., & Stroh, M. (2021). The promises and pitfalls of machine learning for predicting stock returns. *The Journal of Financial Data Science*.
35. Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The*

Journal of Finance, *20*(4), 587-615.
36. Luo. (2021). *Validation of q5--Factor Model in Taiwan Securities Market*
37. Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, *60*(2), 151-158.
38. Menchero, J., & Lee, J.-H. (2015). Efficiently Combining multiple sources of alpha. *The Journal of Investment Management*, *13*(4), 71-86.
39. Menchero, J., Orr, D., & Wang, J. (2011). The Barra US equity model (USE4), methodology notes. *MSCI Barra*.
40. Newey, W. K. (1987). Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics*, *36*(3), 231-250.
41. Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425-442.
42. Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management*, *18*(2), 7-19.
43. Soliman, M. T. (2008). The use of DuPont analysis by market participants. *The Accounting Review*, *83*(3), 823-853.
44. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, *15*(1), 1929-1958.
45. Tobek, O., & Hronec, M. (2021). Does it pay to follow anomalies research? machine learning approach with international evidence. *Journal of Financial Markets*, *56*, 100588.
46. Wu, C.-w. (2020). Smart Beta Trading Strategies With Deep Learning. *Taiwan Academy of Management Journal*, *20*(2), 77-110.
47. Zong, X. (2020). Mlp, cnn, lstm and hybrid svm for stock index forecasting task to indu and ftse100. *Available at SSRN 3644034*.

Appendix

Appendix A. Raw Characteristics

We improved the feature construction method based on Hou et al. (2020) for constructing raw characteristics data suitable for the Taiwan stock market. Due to the inclusion of semi-annual reports before January 2007 and the subsequent transition to quarterly reports, which poses challenges in handling, we restricted our analysis to data from January 2007 to December 2019. There are six categories with a total of 51 raw characteristics

Table A-1 Overview of Categories

Category (Count)	Raw Characteristic
Momentum (8)	Sue, Abr, Mom1m, Mom6m, Mom12m, Depr, Nincr, Chtx
Value (10)	Bm, Bm_ia, Cfp, Ep, Sp, Lev, Cash, Sgr, Cashdebt, Dy
Profitability (7)	Pm, Roa, Rna, Op, Ato, Chpm, Roe
Investment (10)	Noa, Lgr, Chcsho, Agr, Ni, Grltnoa, Pctacc, Gma, Cinvest, Acc
Frictions (10)	Me, Maxret, Std_dolvol, Beta, Rvar_capm, Dolvol, Rvar_mean, Std_turn, Turn, Ill
Intangible (6)	Rdm, Adm, Herf, Rd_sale, Alm, Hire

<u>Momentum</u>:

Sue:

Proposed by Foster et al. (1984). Sue denotes Standardized Unexpected Earnings is quantified as the fluctuation in split-adjusted quarterly earnings per share from its value four quarters ago, and normalized by the standard deviation of this variation in quarterly earnings over the preceding eight quarters.

$$Sue = \frac{Current\ Quarter\ EBT - EBT\ of\ the\ Same\ Quarter\ Four\ Quarters\ Ago}{Stdev(EBT\ over\ the\ past\ eight\ quarters)}$$

Abr:

Proposed by Chan et al. (1996). We compute the cumulative abnormal stock return (Abr) spanning the timeframe proximate to the latest quarterly earnings announcement date. Abr is the closing price on the next financial statement announcement date divided by the closing price on the current month's financial statement announcement date minus one.

Mom1m, Mom6m, Mom12m:

Accumulated past return over the given period. Mom1m = $\sum_{t=1}^{T} r_{i,t}$, where $r_{i,t}$ is daily return and T is the sum of returns over the past 1 month. Mom6m = $\sum_{m=1}^{6} r_{i,m}$ and Mom12m = $\sum_{m=1}^{12} r_{i,m}$, where $r_{i,m}$ is monthly return.

Depr:

Depreciation is divided by PP&E (Property, Plant, and Equipment). Thus, Depr = Depreciation-based Fixed Assets Growth Rate.

Nincr:

Number of Earnings Increases. Based on the study by Green et al. (2013). We measure the number of quarters in the past eight quarters where the quarter-over-quarter growth rate of pre-tax profit is higher than the growth rate in the same quarter of the previous year.

Chtx:

Change in Tax Expense. We calculate Chtx as follows:

$$Chtx = Tax\ Rate(A)_q - Tax\ Rate(A)_{q-4}$$

Here q is the quarter and *Tax Rate(A)* can be obtained from TEJ PRO.

<u>Value</u>:

Bm:

Book to market ratio. We use the reciprocal of Price-to-Book Ratio (P/B Ratio) – TEJ obtained from TEJ PRO to estimate Bm.

Bm_ia:

Industry-adjusted book-to-market ratio. Company's Book-to-Market Ratio minus the Industry Average Book-to-Market Ratio. Evaluating a company's valuation metrics with its industry peers enables an assessment of the company's value and whether it is relatively undervalued or overvalued compared to comparable firms in the same sector.

Cfp:

Quarterly cash flow-to-price is calculated as the cash flow per share divided by the month-end closing price.’

Ep:

Earnings-to-price is represented by the indicator ‘EPS-TEJ’.

Sp:

Sales-to-price is represented by the indicator ‘Revenue Per Share’.

Lev:

Leverage. We initially used the indicator 'Financial Leverage', but most of the data was filled with NaN values. Therefore, we calculated Lev using the formula Lev = EBIT/EBT (Earnings Before Tax - Interest Expense Rate (B) * Net Operating Revenue).

Cash:

Cash and cash equivalents.

Sgr:

Sales growth.

Cashdebt:

Cash-to-debt is calculated as the 'Cash and Cash Equivalents' divided by 'Debt'.

Dy:

Dividend yield is represented by the indicator 'Dividend Yield-TSE'.

<u>Profitability</u>:

Pm:

Profit margin is represented by the indicator 'Net Profit Margin'.

Roa:

Return on assets is represented by the indicator 'ROA-Comprehensive Income'.

Rna

Rna is the return on net operating assets, as proposed by Soliman (2008), and is calculated as (rolling four quarters of 'Operating Income')/('Current Assets' + 'Property, Plant, and Equipment (PP&E)').

Op:

Operating profitability.

Ato:

Asset turnover.

Chpm:

Change in Profit Margin. We calculate *Chpm* as follows:

$$Chpm = \frac{Net\ Profit\ Margin_q}{Net\ Profit\ Margin_{q-1}} - 1$$

Here q is the quarter and *Net Profit Margin* can be obtained from TEJ PRO.

Roe:

Return on equity is represented by the indicator 'ROE-Comprehensive Income'.

Investment:

Noa:

Net Operating Asset is represented by the summation of 'Current Assets' and 'Property, Plant, and Equipment (PP&E)'.

Lgr:

Growth in Long-Term Debt.

Chcsho:

Change in shares outstanding.

Agr:

Asset growth rate.

Ni:

The net stock issue is represented by the indicator 'Cash Capital Increase (Decrease)-CFF'.

Grltnoa:

Growth in Long-Term Net Operating Assets. We calculate Grltnoa as follows:

$$\frac{Change\ of\ Non\ Current\ Assets}{Total\ Assets}$$

$$Change\ of\ Non\ Current\ Assets = \frac{Non\ Current\ Assets_q - Non\ Current\ Assets_{q-1}}{Non\ Current\ Assets_q}$$

Here q is the quarter.

Pctacc:

Percent Accruals. We calculate Pctacc as follows:

$$Pctacc = \frac{Net\ Income - Cash\ Flow\ from\ Operations}{Total\ Assets}$$

Gma:

Gross Profitability is represented by the indicator ‘Gross Margin’.

Cinvest:

Corporate Investment. We measure Cinvest as Cash flows from investing activities / Net sales.

Acc:

Working Capital Accruals. We measure Acc as changes in the working capital (Current Asset-Current Liability) ratio.

Frictions:

Me:

Market value.

Maxret:

Maximum daily return in a month.

Std_dolvol:

The standard deviation of dollar trading volume(3 months).

Beta:

Market Beta(3 months).

Rvar_capm:

CAPM Residual Variance (3 months).

Dolvol:

Dollar Trading Volume.

Rvar_mean:

Return Variance (3 months).

Std_turn:

The standard deviation of Share Turnover (3 months).

Turn:

Shares Turnover.

Ill:

Illiquidity (3 months) is represented by the indicator ‘Quick Ratio’.

Intangible:

Rdm:

R&D to Market Capitalization.

Adm:

Advertising Expense-to-market.

Herf:

Following the methodology delineated by Hou and Robinson (2006), we harness the Herfindahl index, a measure conceived to evaluate the degree of rivalry across firms, as a quantification tool to capture a company's extent of industry concentration. This study categorizes individual stocks based on the ‘TSE Industry Classification’. The calculation method is as follows:

$$Herf = \frac{{MV_{i,t}}^2}{\sum {MV_{i,t}}^2}$$

Here $MV_{i,t}$ is the market value of stock *i* in industry *t*.

Rd_sale:

R&D to Sales.

Alm:

Asset Liquidity. We calculate Alm as follows:

$$\frac{Cash + 0.75 * Noncash\ Current\ Assets + 0.5 * PP\&E}{Total\ Assets}$$

Here $Noncash\ Current\ Assets = Current\ Asset - Cash$.

Hire:

Employee Growth Rate.

Because there are numerous research variables, there is a concern about the potential occurrence of multicollinearity among them, which could adversely impact the estimation results. Therefore, assessing the Variance Inflation Factor (VIF) is necessary to examine the extent of multicollinearity in the dataset.

Table A-2 VIF for Raw Characteristics

Feature	VIF	Feature	VIF	Feature	VIF
chpm	1.00	beta	1.30	mom1m	1.99
acc	1.01	rvar_capm	1.32	maxret	2.03
lev	1.02	agr	1.34	cashdebt	2.07
lgr	1.02	dy	1.35	ato	2.14
ep	1.02	rdm	1.48	rna	2.28
chtx	1.02	rd_sale	1.54	sp	2.51
hire	1.02	adm	1.55	cash	3.18
cinvest	1.03	alm	1.62	noa	3.41
ni	1.04	pm	1.64	std_turn	3.62
chcsho	1.05	herf	1.64	turn	4.01
sgr	1.06	gma	1.66	roe	4.50
nincr	1.07	rvar_mean	1.80	roa	4.93
grltnoa	1.07	mom6m	1.86	me	5.25
depr	1.08	mom12m	1.86	std_dolvol	5.42
cfp	1.09	ill	1.90	bm_ia	5.66
sue	1.10	abr	1.95	bm	6.11
pctacc	1.14	op	1.96	dolvol	8.30

Variable with yellow background represents moderate multicollinearity.

Appendix B. Optimization Details of Deep Learning Model Parameters

To avoid the common pitfalls in machine learning and deep learning raised by Cao (2023), we performed parameter adjustments and optimizations on the model.

Table B-1 Optimization Parameters

Parameter	Setting
Model Network Layers	[64, 16, 4]
Dropout Layer	Dropout Rate =0.5
Early Stopping Epochs	20
Batch Size	16
Learning Rate	0.002

Appendix C. Risk Model Construction

i. Cross-Sectional Solution of Factor Return

To tackle the problem of factor collinearity, we employed the regression solution introduced by Menchero and Lee (2015). Through a weighted and constrained least squares regression, we derive the stock weight matrix for the pure factor investment portfolio. We can compute the factor return once we have the stock weight matrix. The matrix calculation is shown as follows:

$$\Omega = \mathrm{R}(\mathrm{R}^{\mathrm{T}}\,\mathrm{X}^{\mathrm{T}}\,\mathrm{V}\,\mathrm{X}\,\mathrm{R})^{-1}\mathrm{R}^{\mathrm{T}}\,\mathrm{X}^{\mathrm{T}}\,\mathrm{V} \tag{31}$$

Next, we clearly define each symbol, X is the factor exposure matrix with $\mathrm{N} \times \mathrm{K}$,

$$\mathrm{X} = \begin{bmatrix} 1 & X^{I}_{1,1} & X^{I}_{2,1} & \cdots & X^{I}_{34,1} & X^{D}_{1,1} & \cdots & X^{D}_{4,1} \\ 1 & X^{I}_{1,2} & X^{I}_{2,2} & \cdots & X^{I}_{34,2} & X^{D}_{1,2} & \cdots & X^{D}_{4,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & X^{I}_{1,N} & X^{I}_{2,N} & \cdots & X^{I}_{34,N} & X^{D}_{1,N} & \cdots & X^{D}_{4,N} \end{bmatrix} \tag{32}$$

Here $X^{I}_{i,n}$ represents the exposure of asset *n* to industry factor i and $X^{D}_{d,n}$ represents the exposure of asset *n* to deep factor *d.* According to Menchero et al. (2011), applying appropriate regression weights can reduce the error introduced by individual stock-specific returns on testimatingrisk factor returns. Therefore, our regression weight is set to the square root of asset *n*'s market value, $w_n =$

$\sqrt{mv_n}$. V is the weight matrix with N × N,

$$V = \begin{bmatrix} \frac{\sqrt{mv_1}}{\sum_{n=1}^{N}\sqrt{mv_n}} & 0 & \cdots & 0 \\ 0 & \frac{\sqrt{mv_2}}{\sum_{n=1}^{N}\sqrt{mv_n}} & \cdots & 0 \\ 0 & 0 & \cdots & \frac{\sqrt{mv_N}}{\sum_{n=1}^{N}\sqrt{mv_n}} \end{bmatrix} \tag{33}$$

R is the constrain matrix with K × K-1,

$$R = \begin{bmatrix} & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & -\frac{S_1}{S_I} & -\frac{S_2}{S_I} & \cdots & -\frac{S_{I-1}}{S_I} & 0 & \cdots & 0 \\ & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{34}$$

Here S_i is the sum of the square root of market value of assets in industry *i*. Ω is the stock weight matrix of pure factor portfolio with K × N,

$$\Omega = \begin{bmatrix} \omega_1^M & \omega_2^M & \cdots & \omega_N^M \\ \omega_{1,1}^I & \omega_{1,2}^I & \cdots & \omega_{1,N}^I \\ & & \vdots & \\ \omega_{34,1}^I & \omega_{34,2}^I & \cdots & \omega_{34,N}^I \\ \omega_{1,1}^D & \omega_{1,2}^D & \cdots & \omega_{1,N}^D \\ & & \vdots & \\ \omega_{4,1}^D & \omega_{4,2}^D & \cdots & \omega_{4,N}^D \end{bmatrix} \tag{35}$$

where

ω_n^M: the weight of asset *n* in the market factor

$\omega_{i,n}^I$: the weight of asset *n* in industry factor *i*

$\omega_{d,n}^D$: the weight of asset *n* in deep factor *d*

After obtaining the matrix Ω, all factor returns can be calculated using the following formula:

$$f_k = \Omega \times \mathrm{r} \tag{36}$$

To further control the impact of outliers on the regression model, following Biner et al. (2009), we truncate the u_n of asset beyond threshold $4\widetilde{\sigma_u}$,

$$\widetilde{\sigma_u} = 1.4826 \times med(|u_n - med(u_n)|) \tag{37}$$

$$\xi_n = \begin{cases} sgn(u_n)(|u_n| - 4\widetilde{\sigma_u}) & |u_n| > 4\widetilde{\sigma_u} \\ 0 & otherwise \end{cases} \tag{38}$$

ξ should be removed from the original returns, and we rerun the regression to calculate the factor return $\widetilde{f_k}$ after outlier treatment,

$$r_n - \xi_n = \sum_{k=1}^{K} X_{n,k}\widetilde{f_k} + \widetilde{u_n} \tag{39}$$

ii. Daily Covariance Matrix

Biner et al. (2009) introduced the concept of half-life and applied it to the calculation of the daily covariance matrix. Table C-1 shows the parameter settings for the number of days of the half-life.

$$C_{kl}^{(d)} = cov(f_k, f_l)_t = \frac{\sum_{s=t-h}^{t} \lambda^{t-s}(f_{k,s} - \overline{f_k})(f_{l,s} - \overline{f_l})}{\sum_{s=t-h}^{t} \lambda^{t-s}} \tag{40}$$

h denotes the sample size

λ denotes the exponencial weight, defined as $\lambda = 0.5^{1/\tau}$

τ denotes the half $-$ life

Table C-1 Parameters of Covariance Matrix

Model	Variance Half-life	Correlation Half-life	Sample Size
Short Term	90	180	540
Long Term	250	500	1500

As the portfolio risk prediction in this study is set to use a monthly time frame, after calculating the daily covariance matrices, they need to be adjusted to match the monthly period. Considering time-series correlation, we employ the adjustment method proposed by Newey (1987) as follows:

$$C^{(m)} = 22\left[C^{(d)} + \sum_{\Delta}^{D}\left(1 - \frac{\Delta}{D+1}\right)\left(C_{+\Delta}^{(d)} + C_{-\Delta}^{(d)}\right)\right] \tag{41}$$

here $C^{(d)}$ is the daily covariance matrix introduced in equation (40), $C_{+\Delta}^{(d)}$ is the matrix $C^{(d)}$ lagged by Δ days. D is the maximum lag period, which is set to 15 in this research.

Appendix D. In-Sample Explanatory Power of Risk Model

The risk prediction performance of each factor was quite similar. When using the cap-weighted and reg-weighted weighting methods, the explanatory power of the best and worst-performing factors differed by about 2.5%. This indicates that all factors contributed to the risk prediction results to a certain extent.

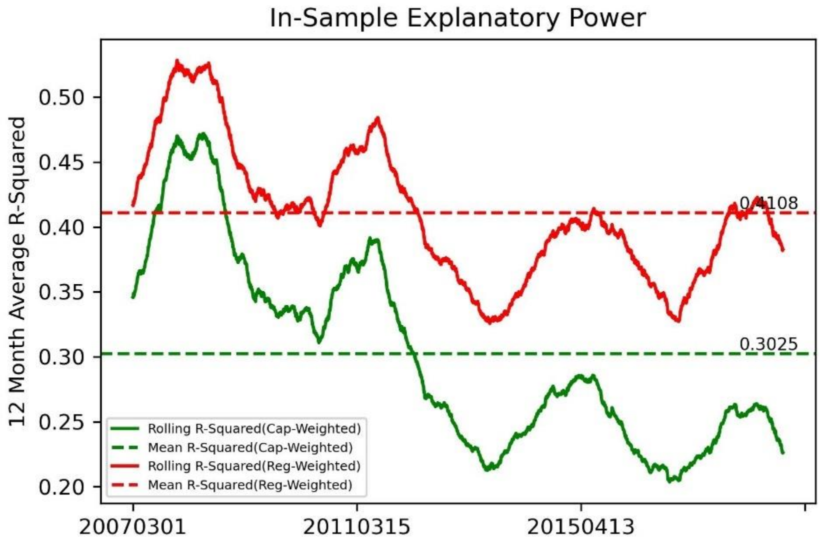

Figure D-1 Explanatory Power of BEST_SHARPE

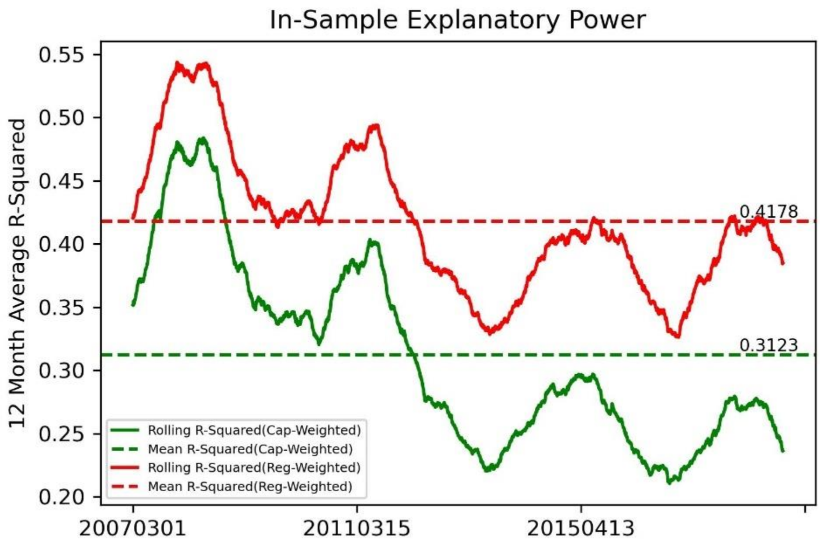

Figure D-2 Explanatory Power of BEST_SR

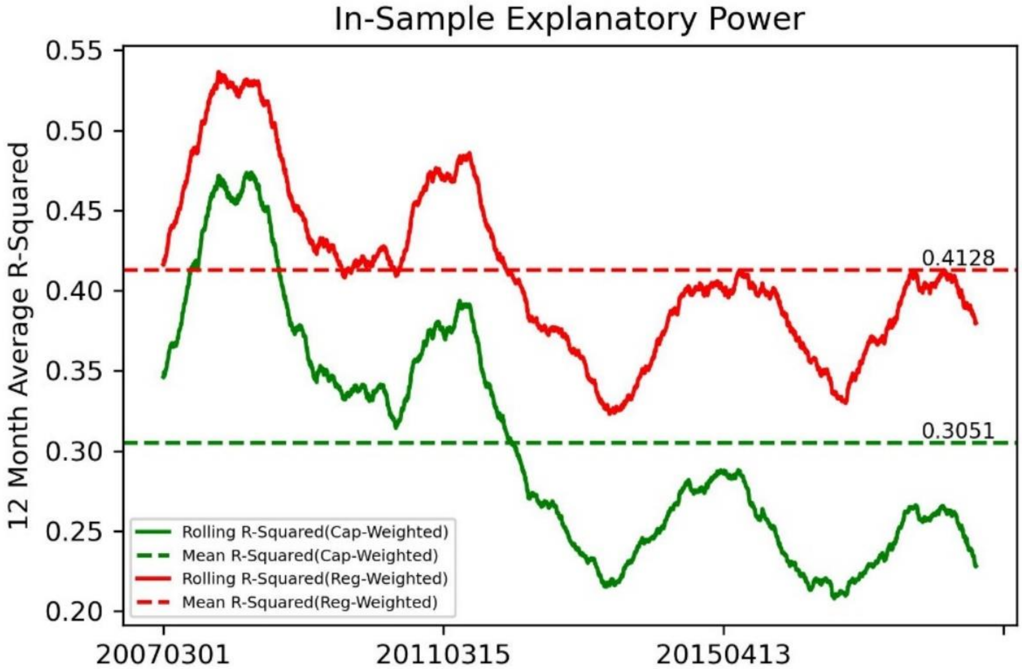

Figure D-3 Explanatory Power of GEN

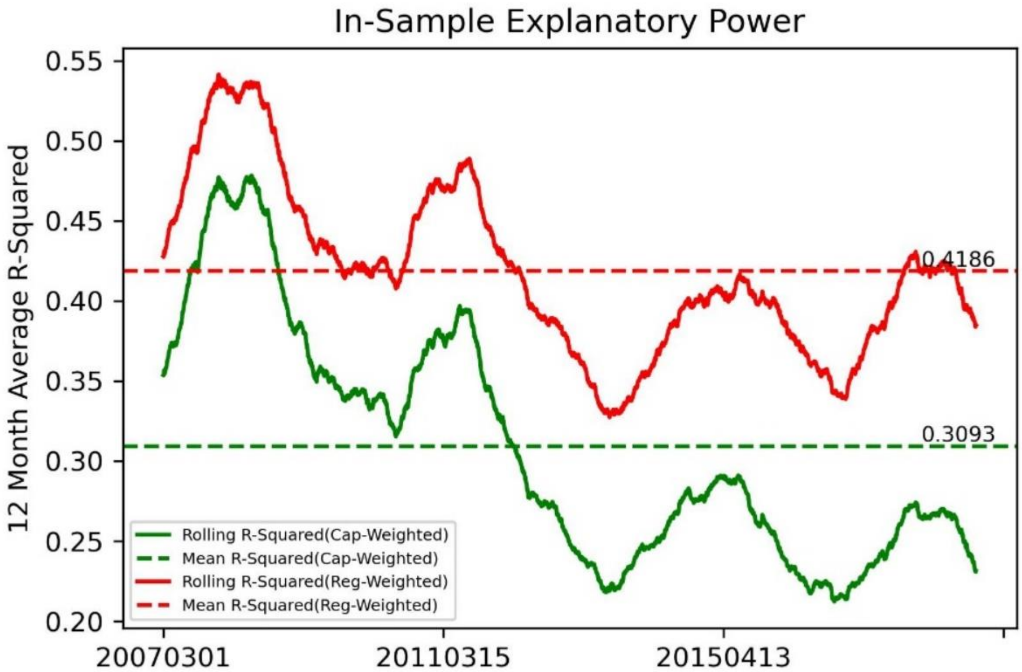

Figure D-4 Explanatory Power of SME

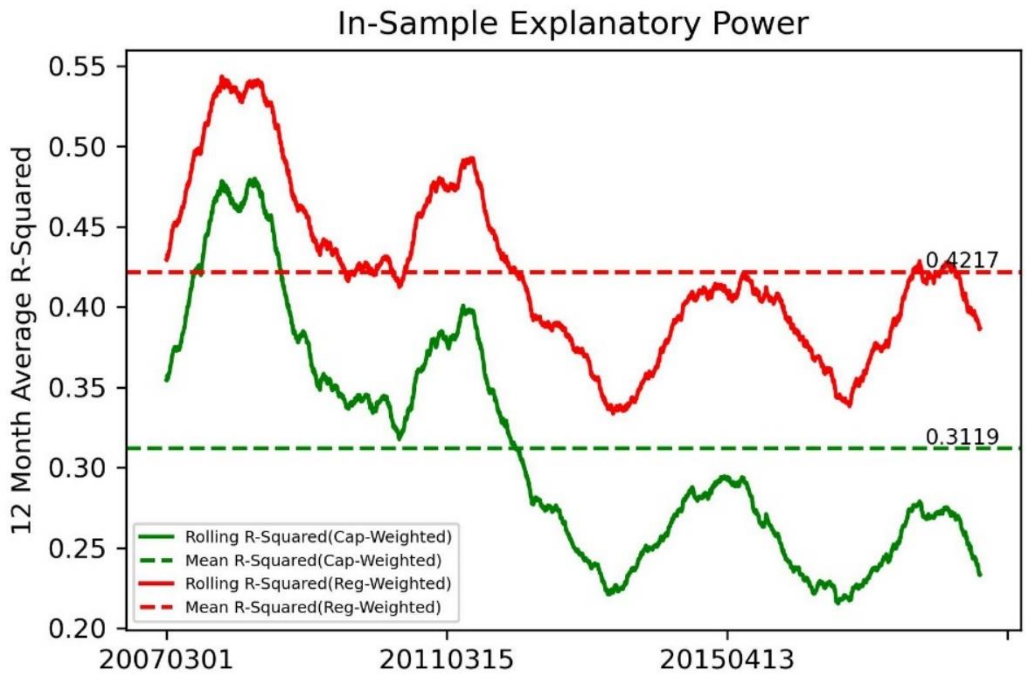

Figure D-5 Explanatory Power of VALUE

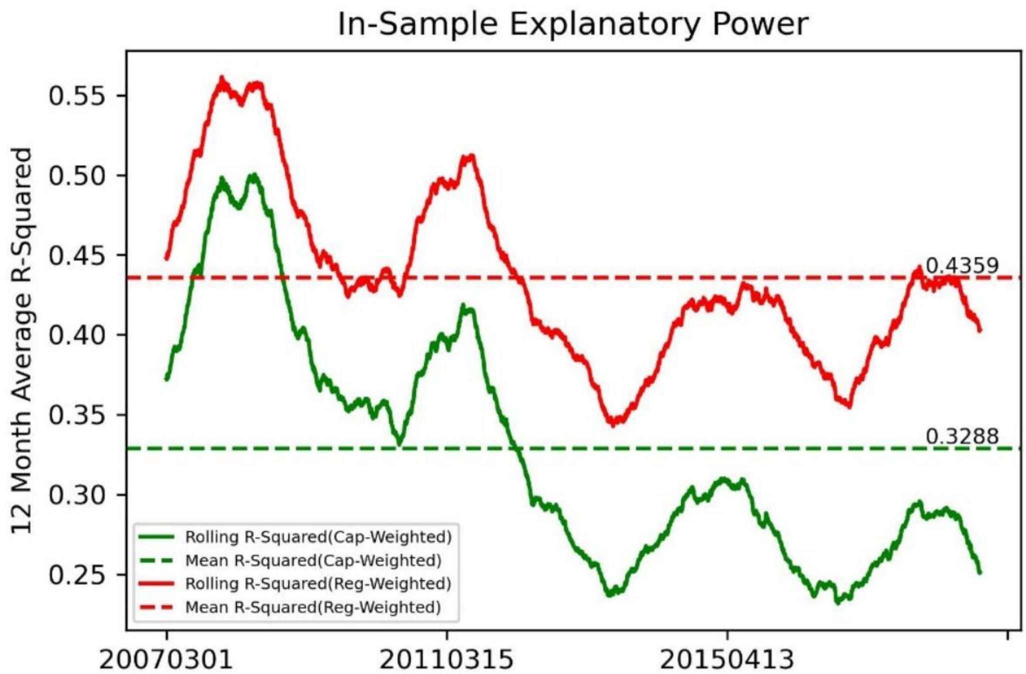

Figure D-6 Explanatory Power of FF