DOI: 10.5281/zenodo.17079646

Factor-based index tracking through the identification of leading stocks

[a*]Mohammad Mokhtari, [b]Sajjad Paykarzade

[a]Department of Management, Faculty of Management and Finance , Khatam University , Tehran, Iran

[b]Department of Management, Faculty of Management and Finance , Khatam University , Tehran, Iran

*Corresponding author: Mohammad.mokhtari507@gmail.com

Published: 08 September2025

Accepted: 11 August 2025

Received: 09 July 2025

Abstract: This paper explores innovative approaches to financial index tracking through the identification of leading stocks using factor models. The primary objective is to introduce a novel, cost-effective method for partial index replication by selecting stocks that exhibit behavior closely aligned with the overall market index. To this end, the proposed model employs Principal Component Analysis (PCA) to estimate latent factors driving asset returns and to identify top-performing stocks. Additionally, a hybrid trading strategy combining momentum and contrarian elements is implemented to enhance portfolio returns relative to the benchmark. The leading-stock-based tracking method not only achieves high accuracy in index replication but also significantly reduces transaction costs and liquidity risk. Simulation results demonstrate that the proposed model outperforms conventional index tracking techniques in terms of minimizing tracking error and generating excess returns. The methodology is applied to data from the Tehran Stock Exchange, encompassing 318 listed companies from the beginning to the end of the year 1400 (2021–2022), and the findings confirm the model's effectiveness in replicating the overall market index. This approach offers portfolio managers a practical tool for constructing low-cost, low-risk portfolios with index-like performance.

Keywords: Index Tracking, Factor Models, Leading Stocks, Momentum and Contrarian Strategies, Principal Component Analysis (PCA)

1. Introduction

This study introduces an innovative and cost-effective approach to index tracking by leveraging leading stocks and a "Follow-the-Leader" algorithm to identify assets with the highest correlation to the benchmark index. By integrating Principal Component Analysis (PCA) and factor models with liquidity and transaction cost considerations, the proposed method enables the construction of precise and economical portfolios that align with the structural characteristics of the Iranian market. This approach represents a meaningful advancement in the development of passive investment strategies. The theoretical foundation of the study draws on classical models such as Modern Portfolio Theory and the Capital Asset Pricing Model (CAPM), which explain the risk-return relationship through the beta coefficient. Despite their limitations, more advanced frameworks like the Fama-French multifactor models have emerged, incorporating variables such as firm size, book-to-market ratio, momentum, profitability, and investment to better capture asset returns under real market conditions. Additionally, the Arbitrage Pricing Theory (APT) offers a more flexible perspective on systematic risk by focusing on macroeconomic factors. In portfolio management, two primary approaches are distinguished: active management, which seeks to outperform the index through continuous market analysis and trading, and passive management, which aims to replicate index performance while minimizing costs and risk. Empirical studies suggest that passive strategies often yield superior results over longer investment horizons. Index tracking, a low-risk method for replicating market performance, typically follows two main strategies: full replication, involving the purchase of all index constituents, and partial replication, which selects a subset of assets and applies optimization algorithms to minimize tracking error. Stratified sampling further refines this process by dividing the index into categories such as industry or firm size and selecting assets proportionally to represent each segment. Advanced versions of these methods employ mathematical optimization to enhance tracking accuracy. Synthetic replication, which substitutes direct asset purchases with derivatives such as futures and swaps, can reduce costs but introduces risks like counterparty default and liquidity constraints. Mathematical models and optimization algorithms play a central role in this domain; heuristic algorithms offer rapid solutions, while metaheuristic techniques—such as genetic algorithms and simulated annealing—approach global optimality. Hybrid algorithms, combining both approaches, improve both the speed and quality of portfolio construction. To simplify portfolio design, PCA reduces data dimensionality and extracts key factors influencing returns, enabling the creation of portfolios that mirror index performance at lower cost. Concepts such as comovement and spillover are also critical in tracker portfolio design: comovement reflects asset correlation with the index, while spillover captures the transmission of shocks across markets. Analyzing these phenomena enhances tracking precision and risk management.

Finally, the study traces the evolution of index tracking methodologies from classical models like the Rudd model and its extensions by Haugen, Baker, and Larsen, to contemporary deep learning algorithms. It demonstrates how the integration of financial theory with modern computational tools can lead to more accurate, cost-efficient, and market-adaptive portfolio solutions—particularly suited to the dynamics of emerging markets.In recent years, studies such as Chuting Sun (2024) have utilized Attention-GRU networks and Conditional Value at Risk (CVaR) models to reduce risk and optimize dynamic portfolios. Other research employing deep learning—particularly neural networks and autoencoders—has demonstrated that these algorithms can more effectively capture complex relationships among index components, thereby enhancing tracking performance. However, several limitations persist. The linear structure of indices, the requirement to use tradable combinations, and theoretical challenges in cointegration models have constrained the free implementation of nonlinear approaches. Corielli (2006) also showed that eliminating low-frequency components from tracking error is only feasible under specific conditions.Overall, index tracking has evolved toward more sophisticated and intelligent modeling techniques, yet it still demands further research in areas such as extreme market volatility, incomplete data, and applicability within less-developed markets.

2. Method

This study aims to track market indices by employing factor models and Principal Component Analysis (PCA) to identify leading stocks. The "Follow-the-Leader" approach selects assets that exhibit the highest correlation with the latent factors underlying the index. After assessing the adequacy of these assets through auxiliary regression, optimal weights are assigned to minimize tracking error. To enhance portfolio returns, two complementary strategies are applied: momentum-based selection (focusing on assets with strong past performance) and contrarian investing (based on the mean-reversion hypothesis). The contrarian strategy, in addition to financial analysis, incorporates behavioral insights into investor psychology and entails higher risk. Successful implementation of these methods requires rigorous analysis, effective risk management, and a deep understanding of market dynamics. The integration of these approaches can improve excess returns and enhance the accuracy of index tracking.

Based on the above explanation, the steps for implementing the contrarian (Follow-the-Loser) strategy are as follows:

1. Selection of Time Period and Data Collection

- Time Period: A specific time frame is selected (e.g., 6 months or 1 year) during which the performance of stocks will be evaluated.
- Data Collection: Closing prices of stocks at the beginning and end of the selected period are gathered to calculate their returns.

2. Calculation of Stock Returns

- Return Formula: The return of each stock over the selected period is calculated using the following formula:

$$r_{it} = \frac{P_{it} - P_{i(t-1)}}{P_{i(t-1)}}$$

Where:

- r_{it} is the return of stock *i* at time *t*
- P_{it} is the closing price of stock *i* at the end of the period
- $P_{i(t-1)}$ is the closing price of stock *i* at the beginning of the period

Stock Ranking

After calculating returns, stocks are ranked from lowest to highest based on their performance. Those with the lowest returns are identified as "losers."

Selection of Loser Stocks

From among the loser stocks, a subset is typically selected as candidates for investment in the upcoming period. This selection may be based on various criteria, such as the largest price decline or the lowest return.

Weight Allocation to Stocks

Weighting Formula: In this approach, stocks that have experienced the greatest decline are generally assigned higher weights. The weight allocation formula can be expressed as:

$$w_i^{adjusted} = \frac{\frac{1}{r_{it}}}{\sum_{j\in L}\frac{1}{r_{jt}}} \,.\, k$$

Where:

- $w_i^{adjusted}$ is the adjusted weight of stock *i*
- r_{it} is the return of stock *i* in the selected period
- (L) is the set of identified loser stocks
- (k) is an adjustment coefficient, typically greater than 1

The final portfolio in this approach consists of stocks with weak past performance, which are assigned higher weights in anticipation of a potential mean-reversion to their historical averages. To maintain effectiveness, the portfolio is periodically reviewed and updated—monthly, quarterly, or annually—so that newly underperforming stocks replace those that have recovered.

Principal Component Analysis (PCA) and the momentum-based "Follow-the-Winner" strategy are both effective tools for portfolio analysis and management, each with its own advantages and limitations.

Advantages of PCA:

- Reduces data dimensionality without losing essential information

- Identifies latent factors influencing asset behavior
- Eliminates multicollinearity for more precise analysis
- Enhances computational speed and efficiency
- Enables data visualization for uncovering investment patterns

Limitations of PCA:

- Potential loss of some information
- Assumes linear relationships among variables
- Difficulty in interpreting principal components
- High sensitivity to data scaling
- Challenges in selecting the optimal number of components

Advantages of the Follow-the-Winner Strategy:

- Capitalizes on positive market trends
- Relies on historical data for confident decision-making
- Simple to implement and effective in identifying successful stocks
- High return potential in bullish markets

In summary, PCA is a powerful tool for reducing data complexity and extracting key factors, while the Follow-the-Winner strategy offers a straightforward and effective means of leveraging market trends. A smart combination of both can significantly enhance portfolio optimization.

This study examines two investment strategies for constructing an index-tracking portfolio: the Follow-the-Winner and Follow-the-Loser approaches.

The Follow-the-Winner strategy is based on the past positive performance of stocks and can generate high returns in bullish markets. Its advantages include ease of implementation and reliance on historical data. However, it carries risks such as buying at peak prices, overlooking intrinsic stock value, and instability in volatile market conditions.

The Follow-the-Loser strategy is grounded in the mean-reversion hypothesis and targets stocks that have experienced significant price declines. Its benefits include purchasing at low prices and capitalizing on market corrections. On the downside, it involves higher risk, uncertainty regarding mean reversion, and potential holding costs.

In summary, both strategies can be effective under specific market conditions. The choice between them depends on thorough analysis, investment objectives, and risk tolerance. In this study, by assigning appropriate weights to each strategy, a portfolio is constructed that not only tracks the index but also generates excess returns.

3. Findings and Results

Computations and Stock Selection The index replication model performs the following computations periodically for each time window:

1. Data Standardization: Input data (stock prices) are first standardized so that each column has a mean of zero and a standard deviation of one. This preprocessing step enhances the model's analytical efficiency.
2. Covariance Matrix and PCA Calculation: The covariance matrix is computed, followed by Principal Component Analysis (PCA) to extract the principal components. This analysis reduces data dimensionality and focuses on components that explain the greatest variance.
3. Stock Selection: Based on PCA results and the discrepancy between standardized and reconstructed data, stocks with the highest and lowest similarity to the index are selected.
4. Weight Optimization: Optimal weights for the selected stocks are calculated using the optimization method described in the previous chapter, aiming to minimize tracking error relative to the market index.
5. Portfolio Return Calculation: Portfolio returns are computed for each time window and compared with market index returns. Cumulative return results are plotted to evaluate model performance and benchmark alignment.

These parameters and analytical steps enable the model to simulate portfolio returns with precision and efficiency, supporting investors in adopting suitable investment strategies.

Baseline Model

Finally, based on the above assumptions, the baseline model is implemented, and its portfolio return and tracking error are evaluated. A summary of the performance of the index-tracking portfolio under review is presented in the Figure 1.

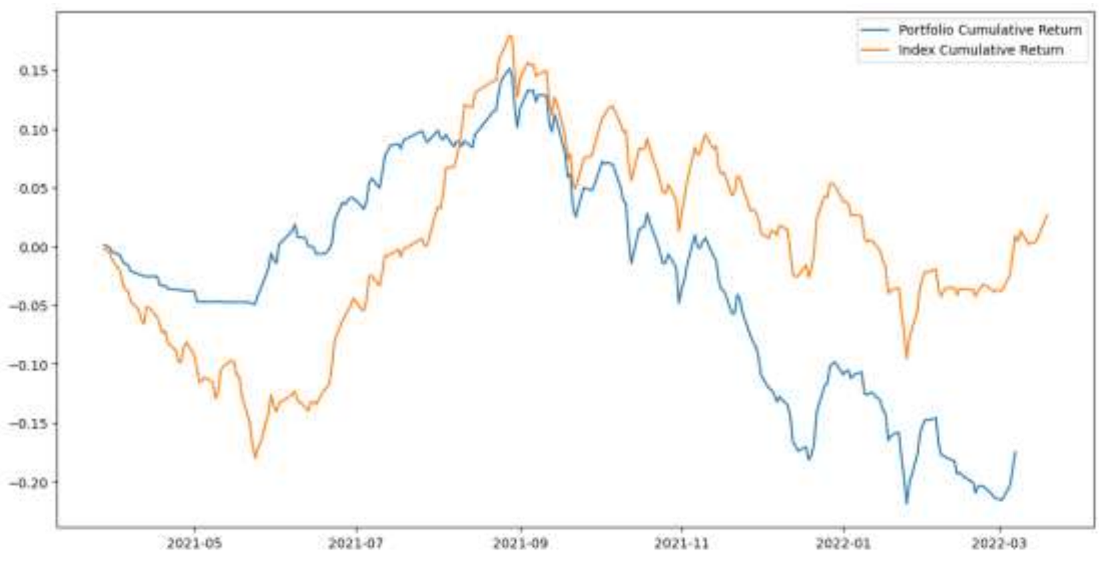

Figure 1. *Portfolio Return of the Index-Tracking Strategy vs. Market Index*

As observed, the index-tracking portfolio has mimicked the benchmark's return with a relatively leveraged behavior. While it experienced greater losses during downward market trends, it also achieved higher returns during bullish phases. Given the prolonged downward trajectory of the index, the final one-year return of the tracking portfolio was –14%, significantly underperforming the benchmark index, which posted a 2% gain. Although the portfolio did not generate excess returns relative to the index, it demonstrated a reasonable ability to replicate the index's return pattern.

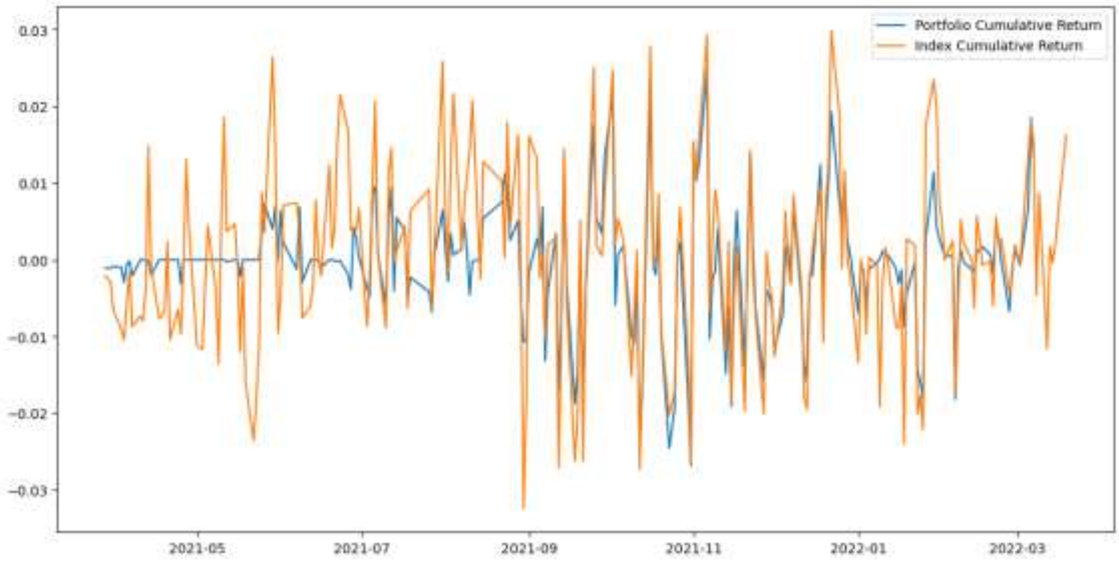

Figure 2. *Replication of Index Trend*

If we examine daily returns instead of cumulative returns, it becomes evident that the baseline portfolio closely follows the trend of the index, although its volatility is lower. This reduced fluctuation may be attributed to a model constraint that limits stock selection to those with a standard deviation below 2%—considered the average market volatility threshold. The tracking error of the index in this model is illustrated in the Figure 3.

Figure 3. *Index Tracking Error*

As observed, the index tracking error has remained below 1% in all but two time windows, with an average error of approximately 0.5%, indicating strong performance in replicating the index trend. In the following section, to enhance model performance, the momentum-based "Follow-the-Winner" strategy will be integrated into the leading-stock tracking framework.

Leading and Momentum Stock Tracking Model

Portfolio performance enhancement is approached by incorporating high-performing stocks (i.e., past winners) as an additive strategy. This method is specifically designed to examine whether selecting and holding stocks with strong historical performance can improve overall portfolio outcomes. This section outlines the theoretical foundations and implementation procedure of the strategy. In efficient markets, prices are expected to fully and immediately reflect new information. However, in practice, markets may be imperfect and influenced by irrational investor behavior. The momentum-based tracking model relies on the hypothesis that a stock's historical performance may signal its future returns. These strategies—commonly referred to as momentum strategies—advise investors to buy stocks with strong past performance and sell those with poor historical returns. The core assumption is that past performance can serve as a predictor of future outcomes. In the leading-stock tracking framework, integration with the momentum strategy is carried out as follows: Each time window defines a historical period during which stock performance is evaluated. Specifically, stocks with the highest cumulative returns during the window are identified as winners. These top-performing stocks are then added to the selected portfolio, with a portion of the portfolio weight allocated to them (e.g., 20%). This allocation is based on the assumption that winner stocks are likely to continue outperforming in the future. The 20% weight is a configurable parameter and can be optimized to improve model performance. The remainder of the portfolio is constructed similarly to the baseline model, with optimal weights assigned to each stock based on its correlation with the index. Optimal weights for both selected stocks and winner stocks are calculated, and portfolio returns are computed separately for strategies with and without the inclusion of winner stocks. This comparison facilitates analysis of the impact of adding momentum stocks on portfolio performance and risk.

A summary of the performance of the momentum-enhanced tracking model is presented in the Figure 4.

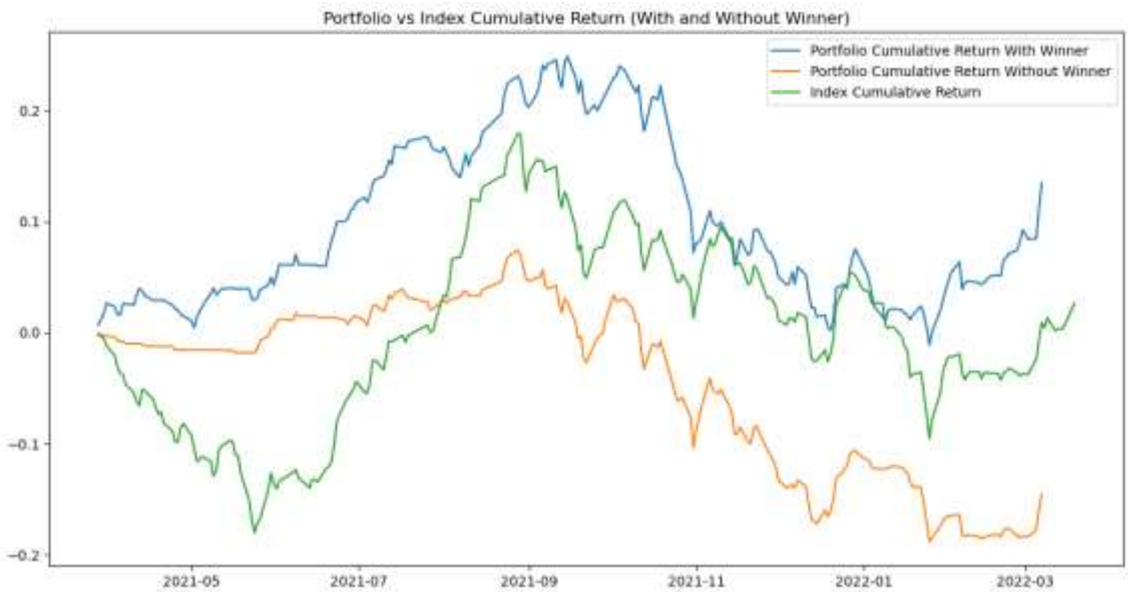

Figure 4. *Portfolio Return of the Index-Tracking Strategy with Integrated Momentum (Follow-the-Winner) Approach*

The results analysis focuses on evaluating the impact of incorporating winner stocks on portfolio performance (cumulative return) and risk (tracking error). These findings are particularly used to assess whether the inclusion of high-performing stocks leads to improved portfolio outcomes and reduced risk. As illustrated in the figure above, the return of the combined leading and winner stock tracking model clearly outperforms the baseline model. Moreover, while effectively replicating the index trend, it also generates excess returns relative to the benchmark.

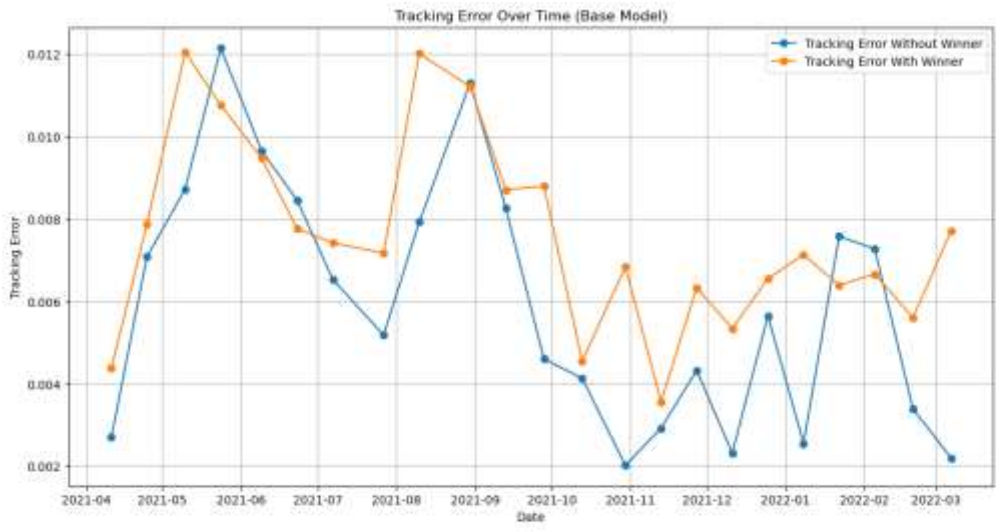

Figure 5. *Tracking Error of the Index-Tracking Portfolio*

Although the enhanced model exhibits a higher tracking error compared to the baseline, it still closely follows the index on average. In a further attempt to improve the model under review, the contrarian strategy—tracking loser stocks with lower returns in the previous time window—was also incorporated. The theoretical foundation of this model is based on the deviation of stock returns from the market average, emphasizing that, over the long term, stocks in the market should converge toward the average market return. Under this assumption, stocks with weaker performance in prior periods may compensate for their lag and deliver higher returns in the current period. The average portfolio returns and index tracking error are presented in the table below for evaluation.

Metrics	Base Model	Follow the Winner	Follow the Loser
Cumulative Return	-0.145579	0.134898	-0.125610
Average Tracking Error	0.005952	0.007577	0.009209

Figure 6. *Portfolio Return and Tracking Error Across Different Models and the Market Index*

Given the negative returns of the contrarian (Follow-the-Loser) model and its relatively high tracking error, further analysis of this strategy has been discontinued.

To evaluate the effectiveness of the momentum-based (Follow-the-Winner) model and determine whether the constructed portfolio accurately reflects the index, the Pearson correlation test is employed. The Pearson correlation coefficient between the index returns and the portfolio returns is 67%, with a p-value of 0.0001—indicating a statistically significant correlation between the two-return series.

For comparative purposes, the Corielli (2006) model—hereafter referred to as the CM model—can be used. This model adopts a relatively similar approach to constructing a factor-based index model. Its key distinction lies in the auxiliary regression and the stock selection process, which incorporates fewer data points and relies on residual correlation testing with the index. Additionally, the weighting methodology in the CM model differs slightly from the approach used in this study.

It is worth noting that, since the PCA model uses data windows extending from the beginning of the evaluation period to the end of each time window, its accuracy improves in later periods. This can be observed through lower volatility and higher tracking error in the initial windows. The Follow-the-Winner model, however, selects more volatile stocks in early periods, which results in greater fluctuations in portfolio returns during those stages.

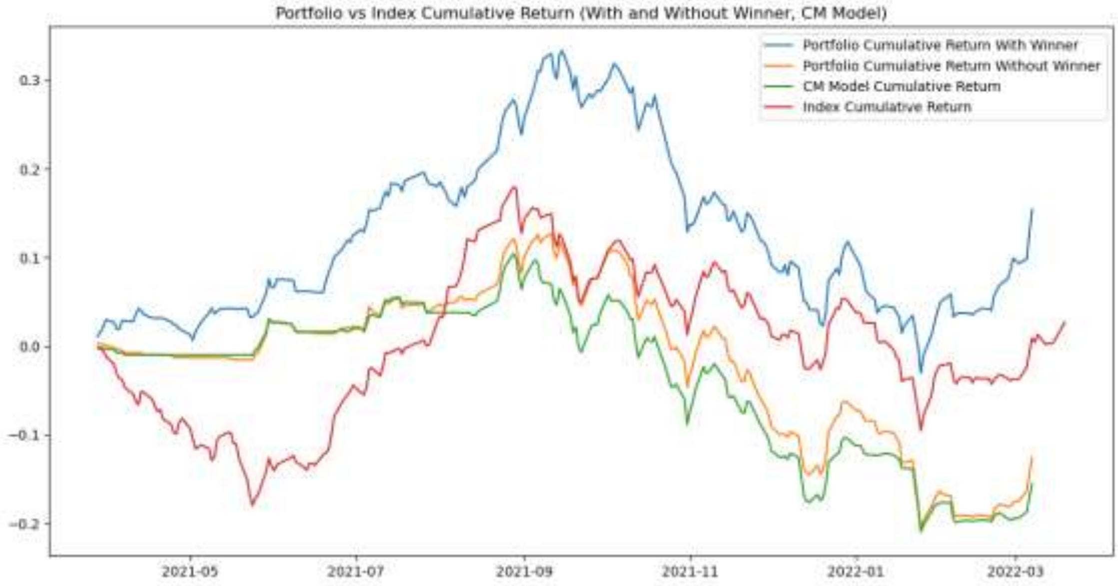

Figure 7. *Comparison of Portfolio Returns Across Different Tracking Models and the Market Index*

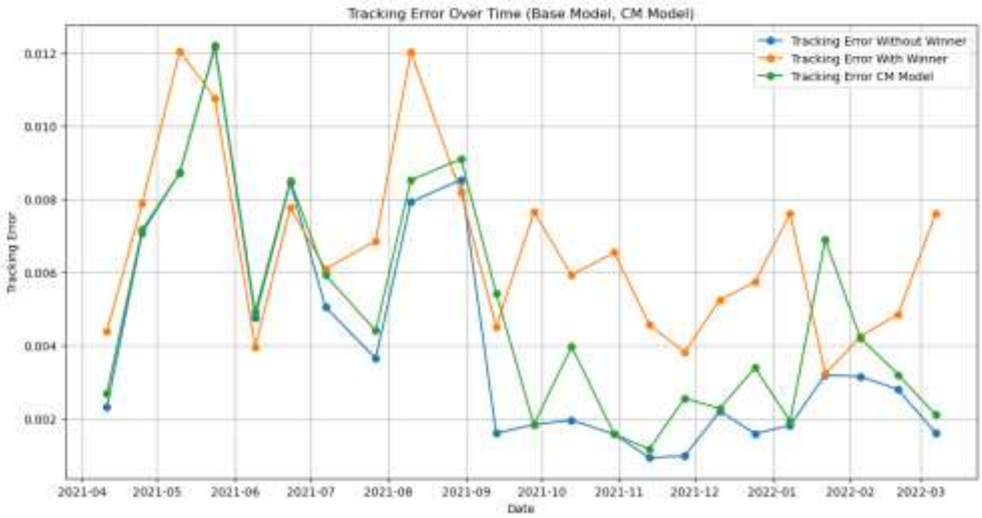

Figure 8. *Comparison of Tracking Error Across Different Index-Tracking Portfolio Models*

The Corielli model exhibits a higher tracking error compared to the baseline model, and its return is lower—though relatively close to that of the baseline.

4. Leading and Contrarian Stock Tracking Model

Portfolio performance enhancement was explored by introducing loser stocks—those with the weakest historical performance—as a new strategic component. This strategy was specifically designed to assess whether selecting and holding underperforming stocks could improve overall portfolio outcomes. According to the Efficient Market Hypothesis (EMH), prices should rapidly and fully reflect all available information. However, in practice, markets may be influenced by irrational investor behavior and overreactions. The contrarian tracking model is based on the hypothesis that stocks with poor past performance may deliver stronger future returns. This assumption stems from the concept of **mean reversion**, which suggests that extremely poor or exceptional performance tends to revert toward the market average over time.

In this model, a historical time window is defined for each evaluation period. Within each window, stocks with the lowest cumulative returns are identified as losers—i.e., the worst performers in the previous period. These loser stocks are then added to the portfolio, with a portion of the portfolio weight allocated to them (e.g., 20%). This allocation is based on the assumption that these stocks may recover from past underperformance and generate higher future returns. The 20% allocation parameter was assigned to loser stocks, but optimization results showed that this adjustment had limited impact on improving portfolio returns. The remaining portion of the portfolio is constructed similarly to the baseline model, with optimal weights assigned based on each stock's correlation with the index.

Finally, optimal weights for both loser stocks and other selected assets are calculated, and portfolio returns are compared with and without the inclusion of loser stocks. This comparison helps evaluate the impact of adding contrarian elements on portfolio performance and risk.

A summary of the contrarian tracking model's performance is presented in the Figure 9.

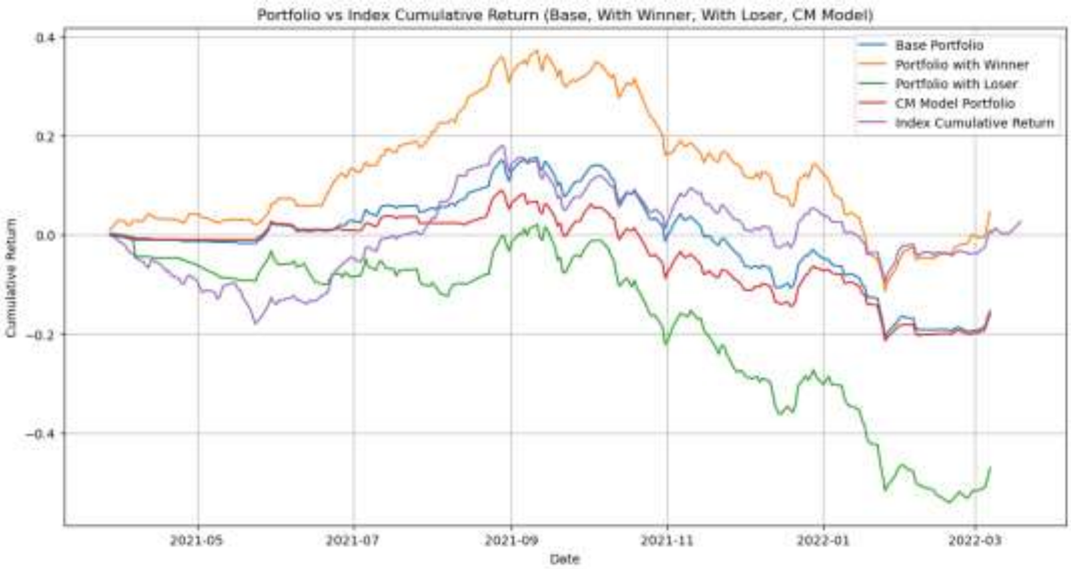

Figure 9. *Comparative Performance of Index-Tracking Portfolios Across Different Models and the Market Index*

This analysis evaluates the impact of incorporating loser stocks on portfolio performance (cumulative return) and risk (tracking error). The results are particularly used to assess whether adding underperforming stocks can enhance portfolio outcomes or reduce risk. As illustrated in the figure, the contrarian (Follow-the-Loser) model has generally produced lower returns compared to the market index. Furthermore, in terms of tracking error, the model demonstrates weaker performance relative to the momentum-based (Follow-the-Winner) strategy.

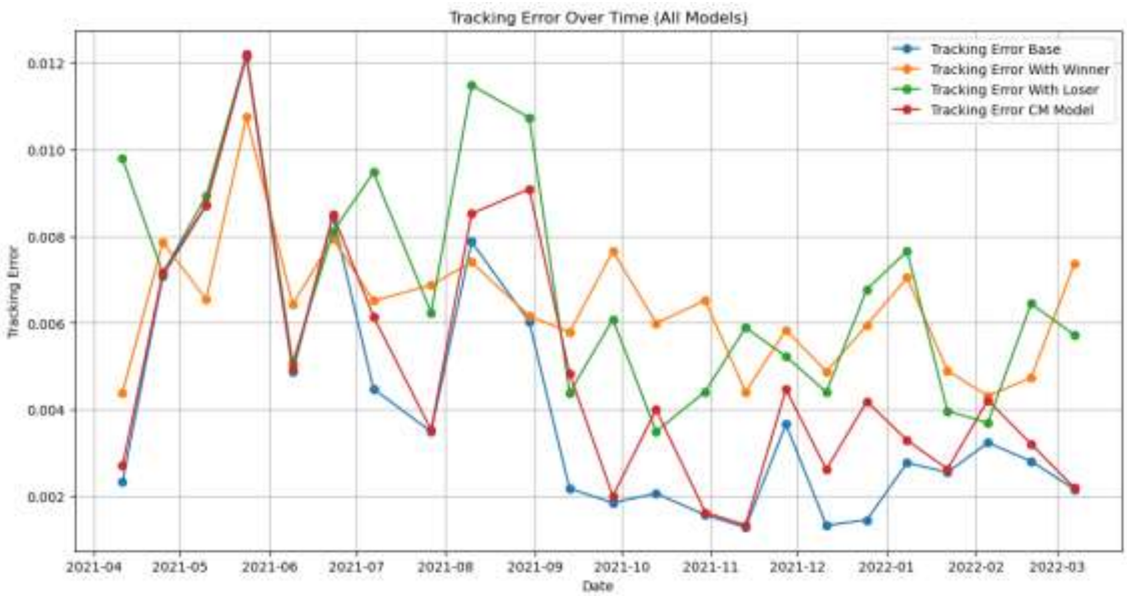

Figure 10. Comparison of Tracking Error Across Index-Tracking Portfolios Using Different Models

It can be observed that the return and tracking error of the contrarian (Follow-the-Loser) model are generally worse than those of the momentum-based (Follow-the-Winner) model and the baseline leading-stock tracking model. Given this weaker performance, further analysis of the contrarian model has been discontinued.

5. Selected Stocks

The number of selected stocks in our model is determined based on the number of key underlying factors. This count is set to ensure that the selected components explain more than 95% of the index variance. Across 30 portfolios constructed over 10-day intervals, a total of 110 distinct stocks were selected. Figure 11 illustrates the return trajectories of the five stocks that accounted for the highest cumulative share of portfolio weights across all evaluated time windows.

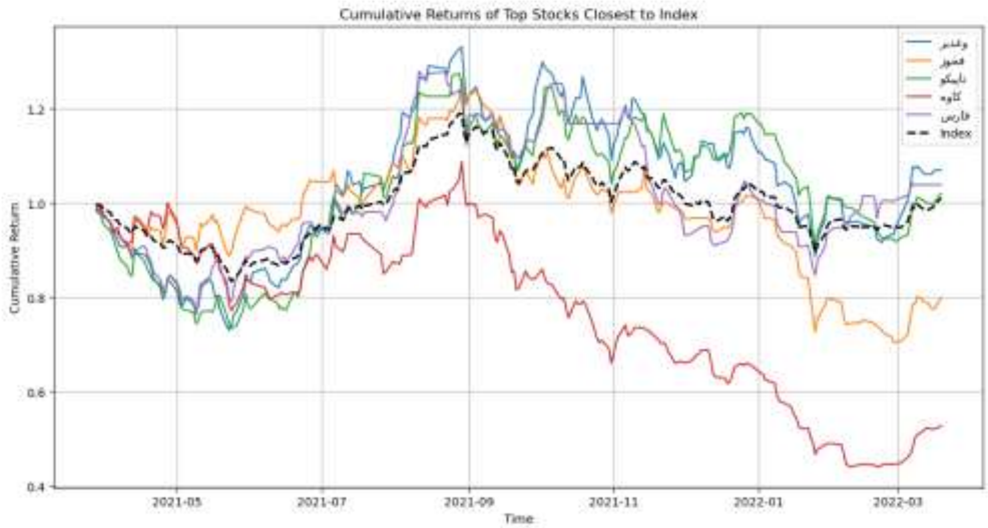

Figure 11. *Top 5 Stocks with the Highest Cumulative Portfolio Weights*

As illustrated in Figure 11, large-cap and index-heavy stocks—such as those in the steel and investment sectors—constituted the largest share of the portfolio throughout the evaluation period. Due to their high correlation with index returns, these stocks contributed to reducing the tracking error in the leading-stock tracking model over time. Collectively, these five stocks accounted for 31% of the total portfolio weight across all examined windows. The selected winner and loser stocks, despite being parametrically assigned 20% of the portfolio weight in this study, do not appear consistently with that proportion. Across the 30 evaluated time windows, a total of 18 winner stocks and 23 loser stocks were selected, each appearing in one or more portfolio periods.

6. Monte Carlo Simulation

In the Monte Carlo simulation phase, we evaluate various portfolio models by randomly selecting time intervals from the return and index datasets. For each interval, the index replication model is executed using the selected data. Based on predefined parameters, the model selects stocks and calculates appropriate weights according to the specific modeling approach. In the Follow-the-Winner model, the top-performing stock from the selected period is added to the portfolio, and new weights are computed accordingly. If any issue arises during weight calculation, default weights are applied and the corresponding error is recorded. For the Corielli model, weights are determined using a method that minimizes index tracking error. Portfolio returns and tracking errors are then computed and stored in their respective lists. These simulation results—comprising cumulative returns and tracking errors—allow us to compare model performance and assess whether the Follow-the-Winner strategy outperforms the CM model. In cases where errors occur during simulation, the iteration is skipped, and default values are added to the results lists to ensure uninterrupted simulation flow. After 1,000 Monte Carlo iterations, the distributions of portfolio returns and tracking errors for both the Corielli and Follow-the-Winner models are analyzed. The results consistently show that the Follow-the-Winner model yields a higher mean return and lower variance. Additionally, its tracking error distribution exhibits a lower average, indicating more stable and accurate index replication.

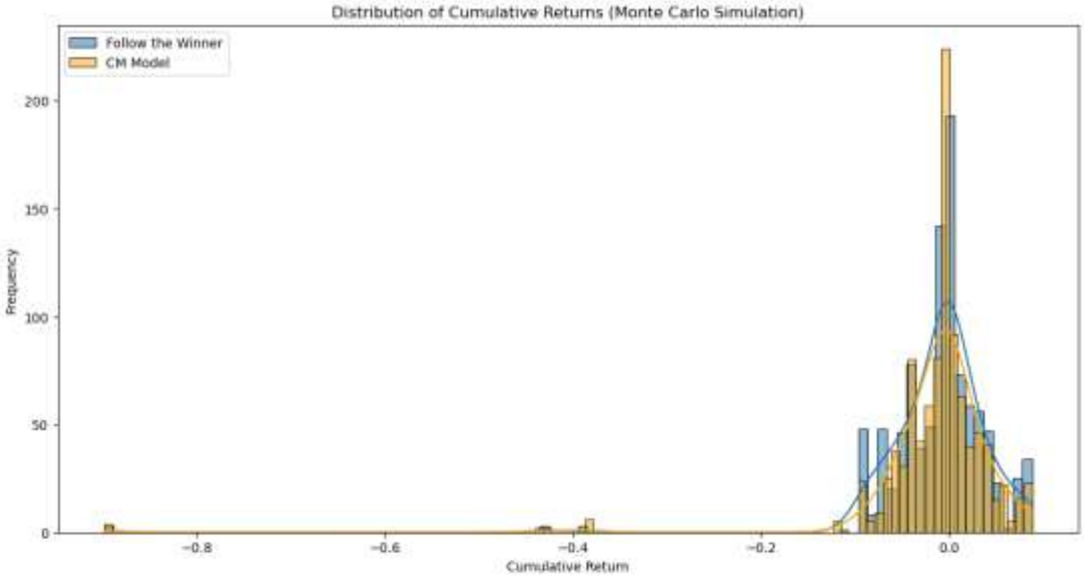

Figure 12. *Return Distribution of the Follow-the-Winner and Corielli (CM) Models*

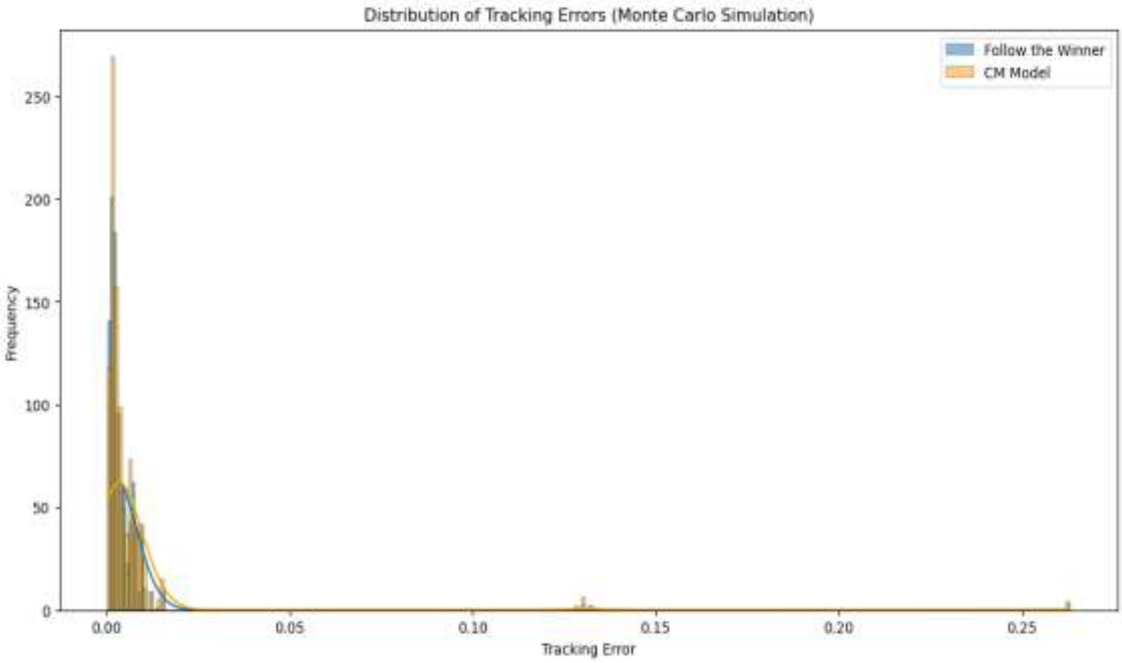

Figure 13. *Distribution of Tracking Error in the Follow-the-Winner and Corielli (CM) Models*

The data consist of daily stock prices from the year (2021–2022), with stocks excluded if they were suspended for more than 150 trading days or added to the index mid-period. Various statistical tests—including normality and autocorrelation—were conducted to assess the data, and their results were analyzed.Subsequently, the baseline index replication model was implemented using Principal Component Analysis (PCA). Key parameters such as the number of principal components and the length of the time window were optimized. Overall, the model successfully simulated index returns with an average tracking error of 0.5%, demonstrating satisfactory performance in replicating market trends.

In the final section, enhancements to the baseline model were explored by incorporating advanced strategies, including leading-stock and momentum-based (Follow-the-Winner) approaches. The results indicate that adding winner stocks to the portfolio led to higher returns compared to the index, although it also increased tracking error. Additionally, Monte Carlo simulations were conducted to evaluate the stability of different models, revealing that the Follow-the-Winner strategy consistently outperformed alternative approaches.

7. Conclusion

Factor models, owing to their unique capabilities in analyzing and forecasting market fluctuations, serve as powerful tools for index tracking. In this study, such models were employed to enhance the accuracy and efficiency of index-tracking portfolios. This chapter was devoted to interpreting the results derived from these models and evaluating their performance in comparison with alternative approaches.

Certain limitations are unavoidable, and uncontrollable factors—often hidden from the researcher—may influence the outcomes of the study. Due to the regulatory structure of Iran's capital market, where price fluctuation limits are imposed on individual stocks and assets but not on the index itself, the data may be subject to bias. Additionally, stocks reopening after suspension may temporarily lack price limits, which introduces further interpretive constraints.

Focusing solely on the Tehran Stock Exchange and the absence of comparative analysis with global financial markets presents a geographical limitation for this research. Moreover, the high degree of correlation among stocks in Iran's market reduces the differentiation between assets, making portfolio weight allocation less sensitive to individual stock selection.

Based on the results obtained from the index-tracking simulations, the proposed model demonstrates high precision in minimizing return deviation from the index. The integration of the baseline model (leading-stock tracking) with momentum and contrarian strategies has generated excess returns relative to the index, offering superior performance for both the portfolio and the investor.

Notably, prior domestic studies have not utilized Principal Component Analysis (PCA) for index tracking. This method, while novel in this context, proves highly suitable and practical—particularly given the leader-centric nature of Iran's capital market and its concentration in a limited number of assets. The findings of this study affirm the validity of the proposed hypothesis.

It is recommended that the tools and methodologies developed in this study be applied to various strategies across different market conditions and time horizons to maximize excess returns over the benchmark index.

The proposed approach can also be extended to other markets or diversified portfolios by incorporating alternative asset classes—such as mutual funds, commodity certificates, or fixed-income securities—to generate unconventional returns relative to the capital market index or other benchmarks, such as commodity or housing indices.

Naturally, the capital market is influenced by a set of underlying variables that shape its trajectory. These methods may be employed to forecast such variables and, consequently, predict the broader trends of the Tehran Stock Exchange index. Further research in this direction is encouraged to deepen understanding and expand the scope of application.

8. References

Bahr-al-Oloom, M. M., Tehrani, R., & Hanafi, F. (2012). Designing a metaheuristic algorithm for optimal index-tracking portfolio selection in the Tehran Stock Exchange. Accounting and Auditing Research Quarterly, 4(13), 22.

Derakhshan, M., Golmakani, H. R., & Hanafi Zadeh, P. (2012). A metaheuristic approach for multi-objective portfolio selection in the Tehran Stock Exchange. International Journal of Industrial Engineering and Production Management, 23(3), 317–331.

Fallahpour, S., Tondnevis, F., & Hashemi, S. M. A. (2015). Optimization of index-tracking portfolios using a stable single-index model based on the top 50 active companies in the Tehran Stock Exchange. Financial Engineering and Securities Management (Portfolio Management), 6(24), 115–134.

Forni, M., & Reichlin, L. (1998). Let's get real: A dynamic factor analytical approach to disaggregated business cycle. Review of Economic Studies, 65, 453–474.

Frino, A., & Gallagher, D. R. (2001). Tracking S&P 500 index funds. The Journal of Portfolio Management, 28(1), 44–55.

Gao, Y., & Li, S. (2018). Index tracking based on LSTM neural network. Procedia Computer Science, 131, 240–246.

Ghalibaf Asl, S. H., Mohammadi, Sh., & Mazaheri Far, P. (2012). Co-movement analysis of stock returns in the Tehran Stock Exchange using the APT model. Management Research in Iran, 16(1), 93–106.

Gilli, M., & Këllezi, E. (2002). The threshold accepting heuristic for index tracking. In Financial Engineering, E-Commerce and Supply Chain (pp. xx–xx). Springer.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223–2273.

Hallin, M., & Liška, R. (2007). Determining the number of factors in the general dynamic factor model. Journal of the American Statistical Association, 102(478), 603–617.

Haugen, R. A., & Baker, N. L. (1990). Dedicated stock portfolios. The Journal of Portfolio Management, 16(4), 17–22.

Hejazi, R., Jafari Saresht, D., & Delshadi, M. (2011). Constructing an enhanced index fund using a genetic algorithm. Tehran Stock Exchange Journal, 4(14), 135–157.

Hodges, S. D. (1976). Problems in the application of portfolio selection models. Omega, 4(6), 699–709.

Huang, N. F., & Guo, J. S. (2009). Partial index tracking using support vector machines. Expert Systems with Applications, 36(2), 2819–2823.

Jiang, P., & Perez, M. F. (2021). Follow the leader: Index tracking with factor models. Journal of Empirical Finance, 64, 337–350.

Jolliffe, I. T. (2002). Principal component analysis. Springer.

Karami, S., & Rastgar, M. A. (2018). Estimating spillover effects of returns and volatility across industries in the Tehran Stock Exchange. Financial Engineering and Securities Management (Portfolio Management), 9(35), 323–342.

Karlow, D. (2012). Comparison and development of methods for index tracking (Doctoral dissertation). Frankfurt School of Finance & Management.

Kim, K. J. (2003). Financial time series forecasting using support vector machines. Neurocomputing, 55(1–2), 307–319.

Kim, S., & Kim, S. (2019). Index tracking through deep latent representation learning. Quantitative Finance, 20(4), 639–652.

Krink, T., Paterlini, S., & Ricciardi, V. (2009). Index tracking: A particle swarm optimization approach. Journal of Economic Dynamics and Control, 33(10), 2003–2019.

Larsen, G. A., & Resnick, B. G. (1998). Empirical insights on indexing: How capitalization, stratification and weighting can affect tracking error. The Journal of Portfolio Management, 25(1), 51–59.

Li, K., & Zhang, L. (2017). Robust index tracking via support vector machines. Journal of Industrial & Management Optimization, 13(1), 429–444.

Li, S., Sun, X., Bao, Y., & Xiong, T. (2011). A heuristic algorithm for index tracking problem. Expert Systems with Applications, 38(5), 5991–5995.

Li, Y., Yang, X., Zhu, S., & Li, D. H. (2014). A hybrid approach for index tracking with practical constraints. Journal of Industrial and Management Optimization, 10(3), 905–927.

Maginn, J. L., Tuttle, D. L., McLeavey, D. W., & Pinto, J. E. (Eds.). (2007). Managing investment portfolios: A dynamic process (Vol. 3). John Wiley & Sons.

Maringer, D., & Oyewumi, O. (2007). Index tracking with constraints using particle swarm optimization. Intelligent Systems in Accounting, Finance and Management, 15(1–2), 57–71.

Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77–91.

Onatski, A. (2009). Testing hypotheses about the number of factors in large factor models. Econometrica, 77(5), 1447–1479.

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. The Review of Economics and Statistics, 92(4), 1004–1016.

Ouyang, H., Zhang, X., & Yan, H. (2019). Index tracking based on deep neural network. Cognitive Systems Research, 57, 107–114.

Parker, J., & Sul, D. (2016). Identification of unknown common factors: Leaders and followers. Journal of Business & Economic Statistics, 34(2), 227–239.

Raei, R., & Telangi, A. (2004). Advanced Investment Management. Tehran: SAMT Publishing.

Razaghi, M., & Hanafi Zadeh, P. (2011). Index-tracking portfolio selection using neural networks. Master's thesis in Financial Engineering, Faculty of Industrial Engineering, University of Science and Culture.

Razaghi, M., & Hanafi Zadeh, P. (2011). Index-tracking portfolio selection using neural networks. Master's thesis in Financial Engineering, Faculty of Industrial Engineering, University of Science and Culture.

Rohweder, H. C. (1998). Implementing stock selection ideas: Does tracking error optimization do any good? The Journal of Portfolio Management, 24(3), 49–59.

Roll, R. (1992). A mean/variance analysis of tracking error. The Journal of Portfolio Management, 18(4), 13–22.

Rudd, A. (1980). Optimal selection of passive portfolios. Financial Management, 9(1), 57–66.

Rudolf, M., Wolter, H. J., & Zimmermann, H. (1999). A linear model for tracking error minimization. Journal of Banking & Finance, 23(1), 85–103.

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

Sharpe, W. F. (1991). The arithmetic of active management. Financial Analysts Journal, 47(1), 7–9.

Shen, Q., Wang, Y., & Ma, C. (2005). Index tracking and statistical factor analysis. Systems Engineering — Theory & Practice, 25(12), 70–78.

Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics, 20, 147–162.

Sun, C., Wu, Q., & Yan, X. (2024). Dynamic CVaR portfolio construction with attention-powered generative factor learning. Journal of Economic Dynamics and Control, 160, 104821.

Taghavi Fard, M. T., Mansouri, T., & Khoshtinat, M. (2007). A metaheuristic algorithm for portfolio selection considering integer constraints. Economic Research Quarterly, 7(4), 49–69.

Wen, F., Yang, X., & Zhou, H. (2018). Deep learning approaches for index tracking: An empirical study in China. Applied Economics Letters, 25

Zare Fazlollahi, F., Sajjadi, J., & Pakizeh, K. (2011). A robust leader–follower approach for portfolio selection. Master's thesis in Financial Engineering, University of Economic Sciences.