Power System Technology
Volume 49, No. 1, 2025
ISSN: 1000-3673

Big Data Analytics in Highly Liquid Exchange-Traded Funds (ETFs) Tracking Performance in China

Avery Miller, Sophia Chen

[1]College of Digital Economics, Nanning University, Guangxi, China

[2]School of Foreign Languages for Business, Guangxi University of Finance and Economics, Guangxi, China

Abstract

As big data technology advances, financial markets have seen a significant rise in data volume and variety, profoundly impacting ETF (Exchange-traded funds) tracking errors. It is crucial to leverage big data technology to analyze new trends and characteristics in ETF tracking errors. The present research studies the tracking performance of 28 high-liquidity stock ETFs in China in the last year. The paper uses tracking difference, daily tracking error(DTE), annual tracking error(ATE), and Panel Linear Regression to evaluate the tracking performance of ETF and the role of the determinants. The study found the underlying benchmark tracking performance of Chinese highly liquid equity ETFs is less efficient than that of more developed region ETFs. The number of stocks included in the underlying index tracked by ETFs has a significant positive impact on the annualized tracking error, while AUM, listing years and daily turnover have a negative impact on the tracking error of ETFs.

Keywords: Big Data, Tracking Performance, Exchange Traded Funds, Tracking Error

1. Introduction

ETF (Exchange-traded funds) refers to open-end funds that copy or track a standardized index by passive management and can be listed and traded [1]. The emergence and development of ETFs is based on Markowitz's portfolio theory [2]. By investing in ETFs, investors can track the underlying index and achieve the investment goal of obtaining benchmark market returns. In addition, the ETF trading method is more flexible. Investors can not only buy and sell ETFs in the secondary market at low cost but also subscribe and redeem them in the primary market. Investors can implement arbitrage operations in the primary and secondary markets [3]. In 1990, Canada's Toronto Stock Exchange (TSE) launched the world's first ETF, Index Participating Shares (TIPs). In 1993, the first ETF in the United States, Standard & Poor's Depository Receipts (SPDRs) was launched [4]. ETFs have developed rapidly around the world in the past three decades. As of June 22, 2022, according to the statistics of ETFGI, there are 9,031 ETFs worldwide.

With the development of big data technology, the volume and variety of data in financial markets have significantly increased [5]. This big data environment has had a profound impact on the tracking errors of ETFs. Big data has not only changed the way investors trade but also enabled fund management companies to more precisely evaluate and adjust their fund portfolios, thereby influencing ETF tracking errors[6]. Through big data analysis, fund managers can monitor market fluctuations in real-time, adjust investment strategies, and reduce tracking errors. Additionally, big data can be used to predict market trends and optimize ETF asset allocation. However, the use of big data also brings new challenges, such as data noise and increased model complexity, which can potentially lead to higher tracking errors [7]. Therefore, in the context of the big data era, leveraging big data technologies to study new trends and characteristics of ETF tracking errors plays a crucial role in further reducing tracking errors and improving performance.

The Shanghai Stock Exchange in China launched its first stock ETF, the SSE 50ETF, in 2005. Over the next 17 years, the number of ETFs in China grew at an unprecedented rate. By July 2022, the Shanghai and Shenzhen stock exchanges had collectively listed a total of 722 ETFs, demonstrating the rapid expansion of this investment vehicle in the Chinese market. Among these, 590 are stock ETF products, marking an increase of 88 from the beginning of the year, indicating the growing interest and demand for ETFs among investors.

2. China's ETF Market

The market value of ETFs listed on the Shanghai Stock Exchange has reached a substantial 1.2 trillion yuan, underscoring the significant role these financial instruments play in China's financial ecosystem. In contrast, the total market value of ETFs listed on the Shenzhen Stock Exchange is comparatively smaller, at 31 million yuan. This disparity highlights the dominant position of the Shanghai Stock Exchange in the ETF market within China. These figures not only reflect the explosive growth of ETFs in China but also emphasize the increasing importance of these funds in the broader financial markets. The data presented in Table 1 below offers a detailed breakdown of the various types of ETFs listed on both the Shanghai and Shenzhen stock exchanges, further illustrating the diverse range of investment opportunities available to market participants. This rapid expansion signals a maturation of China's ETF market, with implications for both domestic and international investors.

Table 1. Category of ETFs listed on the Shanghai and Shenzhen Stock Exchange

Category	Amount	aggregate market value	proportion of amount
Stock ETFS	590	7552 billion	81.72%
Bond ETFS	13	347 billion	1.80%
Commodity ETFS	19	259 billion	2.63%
Monetary ETFS	27	335058 billion	3.74%
Cross-border ETFS	73	1633 billion	10.11%

Data Source: Shanghai Stock Exchange, Shenzhen Stock Exchange

It can be seen from the above data that the number of stock ETFs accounts for 81.72% of all ETFs, which are the main varieties of exchange-traded ETFs and are favored by equity investors. At the same time, the tracking performance of ETFs has attracted more and more attention from scholars and investors. The tracking performance of the underlying index of a stock ETF directly affects the deviation between investor returns and the underlying index return. Tracking performance also affects hedgers' risk exposure. Therefore, tracking performance of stock ETF funds is an important consideration for investors to choose ETFs. At the same time, reducing tracking error is the primary fund management goal of equity ETF fund managers.

The issuance of ETFs in China's securities market is relatively fast, 88 stock ETF funds have been issued and listed from January to July 2022. Whether were the issued stock ETFs in the past five years, especially the highly liquid ETF funds, tracking underlying index efficiently or not? What factors affected the tracking error of this type of fund is a problem worthy of further study. Based on the related literature review, the present research empirically studies the tracking performance and tracking deviation of high-liquidity stock ETFs Shanghai and Shenzhen stock exchanges in China, and then uses the panel linear regression method to explore the influencing factors that affect the tracking performance of high liquidity stock ETFs. In order to provide a reference for investors to make ETF investment decisions, and to provide suggestions for fund managers to reduce the difference between the fund NAV growth and the return of the underlying index.

3. Samples and Research Methodology

3.1 Sample selecting

This study selected 28 stock-type Exchange Traded Funds (ETFs) with high liquidity listed on the Shanghai Stock Exchange or Shenzhen Stock Exchange in China as research samples. To ensure the representativeness of the sample and the reliability of the data, the study employed the following rigorous screening steps.

First, the study identified all ETFs listed on the Shanghai Stock Exchange and Shenzhen Stock Exchange that primarily invest in stocks. During this process, ETFs that target bond markets, money markets, commodity markets, and cross-border markets were excluded to avoid potential biases arising from differences in investment strategies and risk characteristics. Through this screening step, the study obtained an initial pool of 590 stock-type ETFs.Second, the study further excluded index-enhanced funds. Unlike passively managed ETFs, index-enhanced funds aim not merely to replicate index returns but to achieve excess returns through active management strategies. Given the significant differences between these funds and traditional stock-type ETFs, they were excluded from the sample to maintain the study's focus. Following these steps, the study retained 590 stock-type ETFs and further narrowed the sample to those listed before January 1, 2017. This criterion was set to ensure that the selected ETFs had at least five years of historical data, enabling a more comprehensive long-term analysis of their tracking errors. By focusing on ETFs with longer track records, the study aimed to capture trends and patterns that short-term data might not reveal. Additionally, due to the extended investment periods of these funds, their trading characteristics demonstrate relative stability, providing a solid foundation for subsequent research and analysis. After the initial screening, we identified 87 eligible sample funds. However, to further ensure the representativeness of the sample and the accuracy of the research, we raised the selection criteria to include only those ETF funds with a trading volume exceeding RMB 5 billion over the past year (as of July 28, 2022).

Table 2. Research samples: 28 stock ETFs with high liquidity that have been listed for more than five years.

ETF Code	Abbreviation of securities	Benchmark index code	Annualized Tracking Error Threshold [Unit]%	Interval trading volume (including bulk trading) [Transaction Start Date] 1 year before20220728 [Unit] yuan
512880.SH	Securities ETF	512880BI.WI	2.0000	362507662800
512000.SH	Brokerage ETF	512000BI.WI	2.0000	205930600461
159949.SZ	GEM 50ETF	159949BI.WI	2.0000	170952000427
512010.SH	Pharmaceutical ETF	512010BI.WI	2.0000	162882407501
510050.SH	SSE 50ETF	510050BI.WI	2.0000	159320978206
512660.SH	Military ETF	512660BI.WI	2.0000	137626140568
510300.SH	CSI 300 ETF	510300BI.WI	2.0000	103429187934
159915.SZ	GEM ETF Yifangda	159915BI.WI	2.0000	75915873079
512100.SH	CSI 1000 ETF	512100BI.WI	2.0000	59629774938
512070.SH	Securities Insurance ETF	512070BI.WI	2.0000	50291311645
510500.SH	CSI 500 ETF	510500BI.WI	2.0000	45772500406
159928.SZ	Consumer ETF	159928BI.WI	2.0000	40137438862
510330.SH	300 ETF Fund	510330BI.WI	2.0000	27451458301
510880.SH	Dividend ETF	510880BI.WI	2.0000	25725883665
159919.SZ	CSI 300 ETF	159919BI.WI	2.0000	23608147197
510310.SH	CSI 300 ETF E Fund	510310BI.WI	2.0000	16962259649
510210.SH	Shanghai Composite Index ETF	510210BI.WI	2.0000	13593969160
510230.SH	Financial ETF	510230BI.WI	2.0000	13163835730
512500.SH	500 ETF Fund	512500BI.WI	2.0000	12757031501
512810.SH	Defense Military Industry ETF	512810BI.WI	2.0000	9652125953
510580.SH	CSI 500 ETF E Fund	510580BI.WI	2.0000	9289118508
512680.SH	Military ETF Fund	512680BI.WI	2.0000	8903783008
510150.SH	Consumer ETF	510150BI.WI	2.0000	8384390596
159905.SZ	Deep Dividend ETF	159905BI.WI	2.0000	7372091839
159901.SZ	SZSE 100ETF E Fund	159901BI.WI	2.0000	6539304592
159930.SZ	Energy ETF	159930BI.WI	2.0000	6178021048
512600.SH	Required Consumption ETF	512600BI.WI	2.0000	5776342235
510510.SH	CSI 500 ETF Fund	510510BI.WI	2.0000	5197360414

Data Source: Shanghai Stock Exchange, Shenzhen Stock Exchange

After this rigorous selection process, the study ultimately identified 28 stock-type ETFs as the final sample for analysis, which are presented in Table 2. This sample represents the typical characteristics of high-liquidity ETFs in China's ETF market, providing a solid empirical basis for examining the tracking errors of ETFs and their influencing factors in the context of the Chinese financial market.

The basic information, daily closing data, and fund NAV of the sample ETFs used in this research are all sourced from the WIND database. WIND is a leading financial data and analysis tool service provider in mainland China, dedicated to offering comprehensive financial data and information. The WIND database is widely utilized across Chinese financial institutions, with a market share exceeding 90%, and its data quality and coverage are highly regarded within the industry. For this study, data on 28 high-liquidity stock ETFs were extracted from the WIND database, covering a total of 1,334 trading days between January 1, 2017, and June 30, 2022. This time span includes multiple market cycles, providing robust data support for analyzing the long-term tracking performance of ETFs. To ensure the accuracy of data processing and the scientific rigor of the analysis, this research employed Excel and Python software for data organization, processing, and statistical calculations. These tools enabled a detailed empirical analysis of ETF tracking errors, revealing underlying patterns and influencing factors. The chosen methodology aims to enhance the reliability of the research findings, laying a solid foundation for further exploration of ETF tracking performance in the Chinese market.

3.2 Research methodology

In this paper, the research on the tracking performance of sample funds is divided into two parts. In the first part, we calculate the tracking difference and tracking error of high-liquidity stock ETFs between January 1, 2017 and June 30, 2022. In the second part, we analyze and evaluate the possible influencers for the difference in the tracking performance of stock ETFs and the corresponding underlying indexes. The research method is divided into two parts accordingly.

3.2.1 Evaluating tracking performance method

This paper uses three indicators of tracking deviation−−tracking difference, daily tracking error and annual tracking error to evaluate the tracking performance of stock Exchange-Traded Funds (ETFs) against the underlying index.

The tracking difference is the absolute difference between the fund return and the benchmark return. The tracking difference is defined as follows:

$$\mathrm{TD} = \frac{1}{\mathrm{T}}\sum_{t=1}^{\mathrm{T}}|r_{Pt} - r_{Bt}| \qquad (1),$$

the tracking difference according to this definition is also called the absolute mean deviation (Mean Absolute Deviation, MAD), this tracking performance evaluation indicator adopts a linear definition, which is more intuitive and easy to understand. In investment practice, people usually use the linear tracking error definition to evaluate fund managers.

Tracking error, this indicator is the standard deviation tracking difference, which is currently the most common indicator in fund recruitment instructions:

$$\mathrm{DTE} = \sqrt{\frac{\sum_{t=1}^{\mathrm{T}}(e_t - \bar{e})^2}{\mathrm{T}-1}} \qquad (2).$$

Above equation expresses daily tracking error, $e_t = e_P - e_B$ is the sequence of differential returns between the return on ETF portfolio P and benchmark portfolio B on daily basis. $\bar{e}$ is the average value of differential returns, and T represents the number of sample days. Annualized tracking error is calculated by:

$$\mathrm{ATE} = \sqrt{\frac{\sum_{t=1}^{T}(e_t-\bar{e})^2}{T-1}} * \sqrt{N} \quad (3),$$

N represents the number of trading days of the ETF in a year, in this study N is 243.

3.2.2 Track performance affecting factors

Through the analysis of tracking error and tracking difference, we can conclude that the ETFs in the sample exhibited relatively significant tracking errors, meaning there was a noticeable discrepancy between the NAV growth of the ETFs and the returns of their underlying indices. This discrepancy indicates that the ETFs did not fully replicate the performance of their benchmark indices, leading to a mismatch between the actual returns received by investors and their expected returns.

In the next phase of the research, we aim to delve deeper into the factors contributing to this discrepancy. Based on existing research literature and market logic, we have initially identified several key factors that could influence tracking errors. These factors include market volatility, transaction costs, liquidity risk, fund size, management fees, among others. To further validate the significance of these factors and understand their specific impact on tracking errors, this study employs a panel regression analysis method to quantitatively assess these variables[8].

To this end, we have developed the following regression model to analyze the factors influencing tracking errors in China’ s high-liquidity ETFs and their mechanisms of action:

$$\mathrm{ATE_{it}} = \beta_0 + \beta_1 \mathrm{GSi,t} + \beta_2 \mathrm{JJGMi,t} + \beta_3 \mathrm{AGEi,t} + \beta_4 \mathrm{DCJEi,t} + \mathrm{ei,t} \quad (4).$$

Through this model, we can quantify the impact of each variable on tracking errors, thereby identifying which factors have a significant influence on the tracking performance of ETFs. This not only aids in a deeper understanding of the operational mechanisms of ETFs but also provides empirical evidence and strategic insights for fund managers seeking to optimize tracking performance. In the above equation, ATE is the annualized tracking error, GS is the number of stocks included in the index tracked by the ETF, JJGM represents the management scale of the ETF fund, and AGE is the years since listing of the sample ETF. DCJE is the daily turnover of the ETF.

4. Results Achieved

4.1 Research findings on Tracking Error of High-Liquidity stock ETFs

After conducting the calculations, the statistical indicators of tracking performance for the 28 high-liquidity sample ETFs are presented in the following table 3.

Firstly, the table highlights that the tracking difference for each ETF is generally smaller than the daily tracking error, a pattern consistent across all the ETFs listed. The tracking difference represents the absolute deviation between the ETF's return and its benchmark index, while the daily tracking error is the standard deviation of this tracking difference over time. This consistent relationship suggests that, while the ETFs may experience fluctuations in tracking accuracy on a daily basis, the overall tracking difference remains relatively contained. Secondly, a comparison of the empirical daily tracking error results with the target tracking error set in the fund contracts reveals that several ETFs exceed their specified limits. It is observed that several ETFs–specifically 510210.SH, 159930.SZ, 510580.SH, 510880.SH, 512100.SH, 512810.SH, 510150.SH, 159949.SZ, 512680.SH, 510230.SH, 512010.SH, and 512070.SH–exceeded the agreed-upon daily tracking error threshold of 0.05%. These deviations indicate that these ETFs are not tracking their underlying indices as closely as intended on a day-to-day basis, which could be of concern to both investors and fund managers.

The analysis of annualized tracking error provides additional insights. While most ETFs maintain their annual tracking errors within the target limit of 2.00%, there are exceptions, such as 510210.SH and 510580.SH, which exhibit higher annualized tracking errors of 2.44% and 1.70%, respectively. Furthermore, the ETF 510210.SH's annual tracking error exceeds the 2% ceiling, with an empirical result of 0.52%. These findings suggest potential issues in closely tracking the benchmark index over longer periods, possibly due to persistent factors that affect the fund's performance relative to its benchmark.

Lastly, it is worth noting that despite some ETFs exceeding their daily tracking error targets, many still manage to keep their annual tracking errors within the acceptable range. This discrepancy could indicate that the factors driving daily tracking errors, such as short-term market volatility or liquidity issues, may be less impactful over a longer time horizon.

These findings highlight the importance of monitoring ETF tracking performance and suggest that investors should pay close attention to the specific terms outlined in fund contracts, as well as whether the actual performance aligns with these expectations. Rompotis found a mean annual tracking error of 0.63 percent in his research of 50 iShare ETFs domiciled in the US [9]. Johnson et al. found the European ETFs to be generating an annualized tracking error between 0.04–0.21 percent [10]. In this study, the average annual tracking error of the 28 highly liquid ETFs was 0.91 percent. In comparison, the underlying index tracking performance of Chinese most liquid ETFs is larger than that of Western developed countries. The specific reasons need to be further studied.

Table 3. Tracking performance of 28 stock ETFs with high liquidity.

ETF Code	Tracking difference[Unit]%	Daily tracking error[Unit]%		Annual tracking error[Unit]%	
		Target	Empirical Test Results	Target	Empirical Test Results
512880.SH	0.02	0.05	0.03	2.00	0.52
512000.SH	0.02	0.05	0.03	2.00	0.54
159949.SZ	0.02	0.05	0.08	2.00	1.22
512010.SH	0.03	0.05	0.06	2.00	0.88
510050.SH	0.02	0.05	0.05	2.00	0.70
512660.SH	0.03	0.05	0.05	2.00	0.81
510300.SH	0.01	0.05	0.03	2.00	0.40
159915.SZ	0.01	0.05	0.05	2.00	0.77
512100.SH	0.03	0.05	0.09	2.00	1.35
512070.SH	0.02	0.05	0.06	2.00	0.88
510500.SH	0.02	0.05	0.03	2.00	0.40
159928.SZ	0.02	0.05	0.04	2.00	0.65
510330.SH	0.01	0.05	0.03	2.00	0.40
510880.SH	0.05	0.05	0.10	2.00	1.54
159919.SZ	0.01	0.05	0.03	2.00	0.39
510310.SH	0.01	0.05	0.03	2.00	0.41
510210.SH	0.11	0.05	0.16	2.00	2.44
510230.SH	0.02	0.05	0.06	2.00	0.89
512500.SH	0.02	0.05	0.03	2.00	0.42
512810.SH	0.05	0.05	0.08	2.00	1.25
510580.SH	0.05	0.05	0.11	2.00	1.70
512680.SH	0.04	0.05	0.08	2.00	1.17
510150.SH	0.03	0.05	0.08	2.00	1.22
159905.SZ	0.02	0.05	0.05	2.00	0.77
159901.SZ	0.01	0.05	0.03	2.00	0.49
159930.SZ	0.04	0.05	0.12	2.00	1.83
512600.SH	0.03	0.05	0.05	2.00	0.85
510510.SH	0.02	0.05	0.03	2.00	0.51

4.2 Tracking Error influencing factors

The present research uses the panel linear regression model to analyze the possible influencing factors of the annual tracking error, and the results are shown in Table 4.

The regression results indicate that the number of stocks in the index (GS) has a significant positive impact on the annualized tracking error of ETFs. The coefficient is 0.000489, with a T-statistic of 3.787971, and it is significant at the 1% level (P-value of 0.0010). This suggests that when an ETF tracks an index with a larger number of constituent stocks, the annualized tracking error tends to increase. This may be due to the increased complexity of managing a portfolio with more components, making it more difficult for the ETF to accurately replicate the performance of the underlying index, thereby increasing tracking error. The management scale (JJGM) has a negative impact on the annualized tracking error, with a coefficient of -2.60E-11 and a T-statistic of -2.453253, significant at the 5% level (P-value of 0.02222). This result suggests that ETFs with larger management scales tend to have smaller annualized tracking errors. This may be because larger ETFs have more resources to manage and operate effectively, better handling market volatility and operational challenges, thereby reducing tracking error. The years since listing (AGE) also show a significant negative impact on the annualized tracking error, with a coefficient of -0.158153 and a T-statistic of -2.751372, significant at the 5% level (P-value of 0.0114). This indicates that ETFs with longer listing durations tend to have smaller annualized tracking errors. A longer operating history suggests that the fund management team has accumulated more experience, enabling better adjustments and optimization of fund strategies, thus improving tracking accuracy. Daily turnover (DCJE) also has a significant negative impact on the annualized tracking error, with a coefficient of -9.4E-10 and a T-statistic of 2.744394, significant at the 5% level (P-value of 0.0116). This finding suggests that higher daily turnover helps reduce the annualized tracking error of ETFs. Higher turnover indicates better liquidity for the ETF, more efficient price discovery, and reduced price deviations and transaction costs, ultimately leading to lower tracking errors. The constant term (C) in the regression model also shows significance, with a coefficient of 3.741570 and a T-statistic of 3.756553, significant at the 1% level (P-value of 0.0010). This result suggests that there is a baseline level of annualized tracking error even when other variables are not considered, possibly reflecting the influence of market environment, operational risks, or other factors not included in the model. The R-Squared value of the model is 0.688211, indicating that the explanatory variables in the model can account for 68.82% of the variation in annualized tracking error. This suggests that the constructed model has good explanatory power and can accurately capture the key factors influencing annualized tracking error.

The regression analysis reveals that the number of stocks in the index, management scale, years since listing, and daily turnover significantly influence the annualized tracking error of ETFs. Specifically, the more constituent stocks, the higher the tracking error; conversely, larger management scale, longer listing duration, and higher daily turnover are associated with lower tracking errors. These findings provide important empirical evidence for fund managers, helping them to optimize ETF management strategies to improve tracking accuracy.

Table 4. Determinants of Tracking Errors.

Variable	Coefficient	Std.Error	T-Statisic	Prob
GS	0.000489	0.000129	3.787971	0.0010
JJGM	-2.60E-11	1.06E-11	-2.453253	0.02222
AGE	-0.158153	0.057481	-2.751372	0.0114
DCJE	-9.4E-10	3.43E-10	2.744394	0.0116
C	3.741570	0.996012	3.756553	0.0010
R-SQUARED	0.688211			

5. Conclusion

This study investigated the tracking performance of 28 high-liquidity stock ETFs listed on the Shanghai and Shenzhen Stock Exchanges, focusing on ETFs with a trading volume exceeding RMB 5 billion in the most recent year. The analysis revealed that a significant portion of these ETFs exhibited tracking errors that surpassed the limits specified in their fund contracts, particularly in terms of daily tracking error. Specifically, 12 of the ETFs underperformed the upper bound of daily tracking error as stipulated in their respective contracts, while only one ETF exceeded the annualized tracking error threshold.

Moreover, the findings indicated that the tracking performance of Chinese highly liquid equity ETFs is generally less efficient compared to their counterparts in more developed markets such as the United States and Europe. The average annual tracking error for the Chinese ETFs studied was notably higher than those observed in Western markets, underscoring a performance gap that merits further exploration.

The panel linear regression analysis conducted in this study identified several key factors that significantly influence the tracking error of these ETFs. The number of stocks included in the underlying index tracked by the ETFs was found to have a positive correlation with the annualized tracking error. This suggests that as the complexity of the index increases, so does the difficulty for the ETF to accurately replicate its performance, leading to higher tracking errors. Conversely, the regression results showed that larger assets under management (AUM), longer listing durations, and higher daily turnover are associated with lower tracking errors. This implies that ETFs with greater resources, more experienced management, and higher liquidity are better equipped to minimize deviations from their benchmark indices.

Overall, this study highlights the importance of continuous monitoring and optimization of ETF tracking performance. For fund managers, understanding and addressing the factors that contribute to tracking errors is crucial for improving fund efficiency and delivering returns that closely mirror the underlying indices. For investors, these findings underscore the need to carefully consider tracking performance when selecting ETFs, as even small deviations can have significant impacts on investment outcomes over time.

Further research is recommended to explore the underlying causes of the performance gap between Chinese ETFs and those in more developed markets. Such studies could provide deeper insights into the operational and market-specific challenges faced by Chinese ETFs, thereby offering strategies for enhancing their tracking accuracy and overall market competitiveness.

References

[1] Converse, N., E. Levy-Yeyati and T. Williams, How ETFs amplify the global financial cycle in emerging markets. The Review of Financial Studies, 2023. 36(9): p. 3423-3462.

[2] Liebi, L.J., The effect of ETFs on financial markets: a literature review. Financial Markets and Portfolio Management, 2020. 34(2): p. 165-178.

[3] Da Costa Neto, A.F., M.C. Klotzle and A.C. Figueiredo Pinto, Investor behavior in ETF markets: A comparative study between the US and emerging markets. International Journal of Emerging Markets, 2019. 14(5): p. 944-966.

[4] Glosten, L., S. Nallareddy and Y. Zou, ETF activity and informational efficiency of underlying securities. Management Science, 2021. 67(1): p. 22-47.

[5] Ryu, J.P. and H. Shin, Big Data Analysis of Financial Product Transaction Trends Using Associated Analysis. Journal of the Korea Convergence Society, 2021. 12(12): p. 49-57.

[6] Zhang J, Peng Z, Zeng Y, et al, Do big data mutual funds outperform? Journal of International Financial Markets, Institutions and Money, 2023. 88: p. 101842.

[7] Baek S, Lee KY, Uctum M, Oh SH, Robo-advisors: Machine learning in trend-following ETF investments. Sustainability, 2020. 12(16): p. 6399.

[8] Asteriou, D. and S.G. Hall, Applied econometrics. 2021: Bloomsbury Publishing.

[9] Rompotis, G.G., Predictable patterns in ETFs' return and tracking error. Studies in Economics and Finance, 2011. 28(1): p. 14-35.

[10] Johnson, B., Bioy, H., Kellett, A. and Davidson, L., On the right track: Measuring tracking efficiency in ETFs. The Journal of Index Investing, 2013. 4(3): p. 35-41.