MISPRICED AND MISUNDERSTOOD?

THE IMPACT OF LIQUIDITY ON TRACKING ERRORS IN EUROPEAN ETFS

MARIA BANG

VIKTORIA JONASON

Bachelor Thesis
Stockholm School of Economics
2025

Mispriced and Misunderstood? The Impact of Liquidity on Tracking Errors in European ETFs

Abstract:

This thesis investigates how liquidity affects tracking errors in European passive ETFs. To examine this relationship, we apply panel regressions with fixed effects and double-clustered standard errors, controlling for return volatility and fund characteristics. Based on a sample of 494 ETFs from 2000 to 2024, we find that lower liquidity – measured by the relative quoted half spread – is strongly associated with higher pricing-based tracking errors, particularly weakening the price alignment between ETF market prices and their Net Asset Values (NAVs). This finding supports earlier evidence while extending it to the European ETF market. In contrast, tracking errors between NAVs and benchmark indices appear less sensitive to liquidity. The results highlight the importance of market liquidity for ETF pricing efficiency and underscore the challenges posed by Europe's fragmented trading landscape.

Keywords:

ETFs, Liquidity, Tracking Errors, Arbitrage, Panel Regression

Authors:

Maria Bang (25736)
Viktoria Jonason (25735)

Tutors:

Ran Guo, Assistant Professor

Examiner:

Ramin Baghai, Associate Professor, Department of Finance

Bachelor Thesis
Bachelor Program in Business and Economics
Stockholm School of Economics

1 Introduction

Imagine an investment product designed to mimic the market but that, at times, seems to move to its own rhythm. Passive Exchange Traded Funds (ETFs) seek to replicate their underlying indices – yet in Europe, this aim is often complicated by frictions not immediately visible to investors. Since their launch in 1993, ETFs have grown from a niche investment alternative to become one of the most popular product offerings in the asset and wealth management space (PwC, 2022). By the end of 2024, global ETF Assets Under Management (AUM) reached a total value of USD 14.6 trillion, where European ETFs accounted for nearly USD 2.3 trillion (PwC, 2025). Despite this growth, the European ETF landscape remains underdeveloped relative to the U.S., as ETFs represent around 40% of the retail investment market – compared to just 10% to 15% in Europe (Kealy et al., 2025). One feature of the European market worth highlighting is its concentration, where a large portion of ETF investments are directed towards only a handful of large, highly liquid funds, which leaves many ETFs with low trading activity and limited investor attention.

This underdevelopment is not merely a function of scale, but also of structure. According to a report by Vanguard from 2024, European ETFs do not only suffer from lower average trading volumes – around 15 times lower than the U.S. ETFs – but also from a fragmented trading landscape where investors must navigate between different execution venues, currencies and fees. This structural fragmentation has been shown to dilute order-book depth, limit liquidity availability, and reduce the efficiency of arbitrage mechanisms (Vanguard, 2024).

ETF liquidity plays an important part in pricing efficiency. In principle, arbitrage by Authorized Participants (APs) ensure that ETF prices remain closely aligned with their Net Asset Values (NAVs). The APs actively facilitate the creation and redemption of ETF shares in the primary market, as well as assembling the underlying ETF portfolios. Through this process, the supply of ETF shares is adjusted, which helps keep the ETF's market price in line with the value of its underlying assets (Blackrock, 2023). However, APs often struggle to maintain price alignment between ETFs and their underlying assets, specifically when liquidity is limited on either side of the trade. Insufficient liquidity can limit APs ability to execute arbitrage profitably and this can result in persistent mispricing, wider bid-ask spreads, and ultimately higher tracking errors – which is the deviation between the ETF's return and that of its benchmark index (Bae and Kim, 2020).

This issue becomes more than just a technical flaw; illiquidity undermines the key mechanisms that allow ETFs to deliver cost-effective index tracking. In this context, one might ask – are European ETFs, despite their design, sometimes *mispriced* and *misunderstood*? When ETFs are illiquid, APs may struggle to trade at favorable prices, either when establishing arbitrage positions or when closing them to realize profits. Low ETF illiquidity could therefore discourage APs from engaging in arbitrage unless potential profits clearly outweigh trading frictions, which leads to wider bid-ask spreads or postponed arbitrage actions. Such delays in arbitrage can result in higher transaction costs and allow pricing deviations to persist longer than they otherwise would in more liquid environments (Bae and Kim, 2020). These challenges are especially pressing in Europe, where the ETF market is fragmented across multiple national exchanges and regulatory regimes. This results in hindered market-maker activity, and increases the cost of executing arbitrage trades. Consequently, fragmentation exerts an indirect, yet potentially substantial influence on worsening illiquidity and impairing price efficiency.

As a response, Euronext, the pan-European stock exchange group, announced in 2025 a proposal to consolidate over 3300 Exchange-Traded Product (ETP) listings from its seven exchanges – Milan, Amsterdam, Paris, Oslo, Brussels, Dublin, and Lisbon – into a single venue. By consolidating listings onto a single platform, Euronext thus aims to concentrate trading activity, thereby enhancing liquidity. Improved liquidity can facilitate more efficient price discovery and decrease the cost of trading ETFs; benefitting both institutional and retail investors. Enhanced liquidity resulting from consolidation can lead to more accurate pricing of ETFs relative to their underlying assets. This accuracy can minimize tracking errors, since ETFs would be better positioned to reflect the performance of their benchmark indices. A unified trading venue can furthermore streamline the arbitrage process, enabling authorized participants to more effectively align ETF prices with their Net Asset Values (NAVs). (Johnson, 2025)

Although the relationship between liquidity and ETF tracking errors has been examined in U.S. markets (e.g. Bae and Kim, 2020; Ackert and Tian, 2008), empirical evidence for European ETFs remains limited. Existing studies that do include European ETFs often rely on small samples, short time frames, and do not focus on liquidity as a central factor (e.g. Tsalikis & Papadopoulos, 2019). As European ETF markets grow in importance, it is crucial to understand the mechanisms that distort their performance; not only for academics, but for practitioners, regulators, and investors alike. This study addresses this gap by investigating how liquidity influences ETF tracking errors in the European market.

This thesis examines the extent to which ETF illiquidity contributes to tracking errors in the European ETF market. We explore this relationship over a 25-year period (2000 to 2024), covering 494 passive ETFs, listed on European exchanges. Our contribution to the literature is threefold. Firstly, we focus exclusively on Europe, where liquidity is affected by unique structural fragmentation across exchanges. Secondly, our dataset spans more than two decades and includes major financial disruptions such as the global financial crisis in 2008 and the COVID-19 pandemic. Lastly, we apply multiple tracking error measures and a robust panel regression framework with double-clustered standard errors to enhance the credibility of the findings. Our hypothesis is that illiquid ETFs exhibit larger tracking errors.

To test this hypothesis, we estimate the tracking errors using three different methods: (i) absolute return differences, (ii) regression-based – and (iii) standard deviation-based tracking errors. ETF liquidity is proxied using the relative quoted half-spread based on end-of-day ask and bid prices. We compute both annual and daily panel regressions and control for factors such as AUM, ETF type, NAV return volatility, as well as index return volatility. Furthermore, fixed-effects and double-clustered standard errors are applied to mitigate time series and cross-sectional dependencies.

We find that the illiquidity's effect on European tracking errors is multidimensional. ETF pricing-based tracking errors (ETF versus NAV, and to some extent, ETF versus Index) are sensitive to market liquidity conditions, whereas index replication-based deviations (NAV versus Index) are relatively insensitive to such frictions. The consistency of findings across regression types and robustness checks provides evidence that ETF illiquidity plays a central role in shaping pricing misalignments in the European ETF market.

This thesis is organized as follows. In Section 2 we will review the existing literature on which our study is based on. In Section 3, fundamental concepts are presented. In Section 4, we explain the data collection process and the research method used to examine our research question, including our liquidity measure and the regression models. In Section 5, the empirical findings are presented. Section 6 investigate the robustness of our model. Lastly, we provide concluding remarks in Section 7.

2 Literature Review

The effect of liquidity on asset returns have previously been widely studied. Earlier work on the topic, such as Amihud and Mendelson (1986) establish a linear relation between expected return and illiquidity, where a larger bid-ask spread leads to a higher required yield on stocks. Further studies provide additional evidence that illiquidity is priced, for instance Pastor and Stambaugh (2003) demonstrate that stocks with higher sensitivities to innovations in market liquidity earn significantly higher returns, and Acharya and Pedersen (2005) develop a liquidity-adjusted CAPM model and find that stocks with higher exposure to different liquidity risks have higher average returns (Acharya and Pedersen, 2005) (Pástor and Stambaugh, 2003). For ETFs, the liquidity effect is amplified due to the dual-market structure in place; ETFs are traded on secondary markets, such as major stock exchanges, but their shares are created and redeemed by authorised participants in the primary market. Liquidity constraints in either market can thus impair the arbitrage mechanism, when the authorised participants are not able to keep prices aligned by trading efficiently in both the fund and the underlying basket of assets. This affects how accurately an ETF can replicate its benchmark index and will widen bid-ask spreads and raise the total cost of ownership (Bae and Kim, 2020).

Holden et al. (2014) provide a comprehensive framework for understanding and measuring market liquidity. They discuss the concept of *commonality in liquidity,* where liquidity levels tend to co-move across assets, which is something that is particularly relevant to ETFs, as the ETF liquidity is closely linked to that of its underlying assets. This concept is further emphasized by Bae and Kim (2020), whose work show that ETF liquidity is closely linked with the liquidity of the assets in the underlying index the fund aims to replicate. The paper by Holden et al. (2014) also discusses how market structure and regulation influence liquidity. The authors posit that liquidity significantly depends on factors such as exchange transparency, tick-size regulations, and the existence of algorithmic or high-frequency trading, with their analysis indicating that more transparency and reduced fragmentation in the market lead to improved liquidity, decreased trading expenses and a strengthened price efficiency. (Holden et al., 2014)

Bae and Kim (2020) present one of the most thorough studies on the implications of ETF liquidity. They investigate the effect of ETF liquidity on ETF tracking errors, returns and volatility in the U.S. ETF market. Their study finds that illiquid ETFs have sizable tracking errors, and that this effect is more pronounced when the underlying assets are less liquid. Furthermore, they find that liquidity risk earns a premium, where this risk contributes approximately 14.2 basis points (0.142%) to the annualized return difference between portfolios that have been categorized by liquidity levels. This means that a positive liquidity premium exists in the U.S. ETF market, implying that illiquid ETFs yield higher expected returns as compensation for elevated liquidity risk. Additionally, the study identifies how illiquidity reduces the effectiveness of arbitrage. When the ETF or its underlying assets are illiquid, the APs face higher costs and risk when executing arbitrage trades to replicate the ETF's underlying index. Consequently, the APs may delay the trades until the ETF's price deviation from its net asset value is significant, or they may demand higher spreads to compensate for the execution risk. In other words, when the cost of executing arbitrage is too high relative to the price deviation, APs may refrain from immediate action. This hesitancy can lead to persistent pricing mismatches and increased trading costs for investors. Our study extends this analysis to the European ETF market, which differs in structure and regulatory context. (Bae and Kim, 2020)

Ackert and Tian (2008) complement Bae and Kim's (2020) view by investigating the role of liquidity and arbitrage in the pricing efficiency of U.S. and country ETFs. Their research compares U.S. ETFs, which are typically priced close to their NAV, with country ETFs – which often experience significant mispricing, and show that there is a non-linear (inverted-U) relationship between liquidity and pricing deviations. At low liquidity levels, increased trading can worsen the mispricing, whereas at higher levels of liquidity more trading improves pricing efficiency and decreases the ETF's deviations from its NAV. Additionally, the authors find that for U.S. ETFs, higher liquidity is generally associated with lower fund premiums. This aligns with the previously mentioned idea that higher liquidity facilitates arbitrage and corrects mispricing. Although this article focuses on international ETFs traded in the U.S., during an earlier and shorter time period (2002-2005), Ackert and Tian's findings still support the core proposition of this thesis; namely that liquidity is a key determinant of ETF price accuracy. (Ackert and Tian, 2008)

While Ackert and Tian's (2008) work highlight how liquidity levels affect the pricing efficiency of ETFs, Ben-David et al. (2018) examine how ETFs themselves can create volatility in the broader market, especially through their impact on the underlying assets that they track. Ben-David et al. find that stocks with higher ETF ownership exhibit significantly higher volatility compared to stocks with lower ETF ownership. Furthermore, the study finds that if the increase in stock volatility caused by ETFs to some extent is non-diversifiable, and it may constitute a form of systematic risk for certain investors. ETF ownership might therefore call for a risk premium, when market participants require higher compensation for bearing this ETF-induced risk. This is also shown in the study, when the stocks with higher ETF ownership tend to earn higher average returns. (Ben-David et al., 2018)

Empirical studies on European ETFs are sparse. Tsalikis and Papadopoulos (2019) examine how well funds replicate their benchmark indices from a sample of 15 American and European ETFs. Using three different methods for estimating tracking errors, the authors find that American ETFs consistently appear to exhibit lower tracking errors than European ETFs. Furthermore, the authors analyze the factors that affect tracking errors and posits that fund size and expense ratios are some of the factors with the highest influence. Although this study examines European ETFs, only a small sample of 10 European funds are used, and the authors do not examine liquidity as one of the factors affecting tracking errors. Our study aims to examine a larger sample of European ETFs, and includes liquidity as a factor affecting tracking errors. (Tsalikis and Papadopoulos, 2019)

In spite of the growing ETF literature, relatively few studies focus particularly on how liquidity affects European ETFs. The European market is further complicated by its complex market structure and fragmentation, with multiple listing venues, different settlement systems and regulatory frameworks, which all contribute to higher trading costs and wider bid-ask spreads. As Vanguard (2024) outlines, fragmentation in the market can dilute the order book depth, which makes it harder for investors and APs to smoothly match buy and sell orders. This results in a reduction in the overall liquidity available to support arbitrage. This structural complexity may thus aggravate the liquidity challenges highlighted in the studies focused on the U.S. market. (Vanguard, 2024)

This structural issue in the European market has gained increased attention in recent years. Johnson (2025) reports that Euronext, the pan-European stock exchange group, is proposing a consolidation of over 3300 exchange-traded product (ETP) listings from its seven exchanges into a single venue. The goal is to concentrate trading activity, thereby enhancing liquidity, improving pricing efficiency and decreasing trading costs. Improved liquidity would also support more effective arbitrage, which will help align ETF prices closer to their NAVs, as well as reducing tracking errors. These developments reinforce the significance of liquidity as a central concern for investors in European ETFs and underscore the importance of investigating its effects in the context of this thesis. (Johnson, 2025)

To conclude, the literature establishes three stylised facts. Firstly, liquidity is priced and directly affects ETF investors via bid-ask spreads, volatility and a liquidity-risk premium. Secondly, the efficiency of arbitrage decreases under illiquidity, which leads to persistent mispricing and larger tracking errors. Thirdly, the European market structure is complex and uniquely fragmented; something that suggests that liquidity costs documented in the U.S. may be even more pronounced in the European market – although systematic evidence is scarce. This thesis contributes to the literature by focusing on the European ETF environment, investigating whether liquidity exerts a similar influence on tracking errors. Thus, an open question remains: *to what extent does liquidity drive tracking errors in European ETFs?* Addressing this gap is the central contribution of the present thesis.

3 Fundamental Concepts

3.1 Exchange Traded Funds

An Exchange Traded Fund (ETF) offers investors a way to pool their money in a fund that holds a collection of assets such as stocks, bonds, or commodities. By holding a basket of securities, an ETF thus reduces the risk associated with investing in individual assets. ETF shares are traded at market prices on stock exchanges; differing from mutual funds, which are priced based on their Net Asset Value (NAV) at the end of the trading day. The ETF price is typically different from the fund's NAV per share, because the ETF's market price fluctuates during the trading day, due to factors such as the underlying prices of the ETF's assets and the demand for the ETF. (U.S. Security and Exchange Commission, 2024)

While both ETFs and Closed-End Funds (CEFs) are traded continuously through an exchange, the two differ as most ETFs are open-ended funds, which means that shares can be issued and the Assets Under Management (AUM) can grow continuously. CEFs on the other hand, is a type of mutual fund that issues a fixed number of shares through an Initial Public Offering (IPO) and these shares are then traded on the secondary market. The fund will hence always have a specific amount of capital invested. (Horton, 2024)

While an active ETF uses a portfolio manager's investment strategies to try and outperform a benchmark – often a specific index – passive ETFs instead try to match the performance of a benchmark without active decision-making by fund managers, which leads to it tending to be lower-cost than active ETFs (McWhinney, 2024). The passive ETF does not attempt to outperform the benchmark but rather mirrors its returns by holding the same assets in similar proportions (Liberto, 2024).

An index-based ETF aims to earn the return of the market, or the subset of the market, that it aims to replicate, less the fees. An index is designed to, as closely as possible, measure the value of a certain market or a segment of that market. Index ETFs normally are among the lowest-cost ETFs, due to their minimal portfolio turnover and very low research costs (Ferri, 2024). Index-based ETFs also allow investors to reach foreign markets at a relatively low transaction cost (Bae and Kim, 2020).

The core risk of an index-based ETF stems from the market risk linked to the index it follows. Apart from this index-related risk, investors may also face additional risk if the ETF does not accurately replicate the performance of its underlying index. This means that discrepancies between the ETF's returns and the index returns, or its NAV returns can present an extra layer of risk for investors. (Bae and Kim, 2020)

Although ETFs are generally traded on major stock exchanges, their shares are created and redeemed in the primary market. Because ETFs trade both in primary and secondary markets, their shares can exhibit two prices simultaneously: the market price, which is determined through exchange trading, and the NAV, which reflects the value of the underlying securities. Under no-arbitrage conditions, ETF returns are expected to track their NAVs on a daily basis. However, several factors can cause discrepancies between the two. (Bae and Kim, 2020).

ETFs carry different underlying assets, and therefore have different fund characteristics. These characteristics affect liquidity, and generally liquidity transformation is higher in ETFs tracking less liquid underlying securities such as bond ETFs (Grill et al., 2018). A common type is the Equity ETF, which covers large or small business or stocks from a specific country (Blackrock, 2021).

4 Data and Methodology

Our analytical approach builds on insights from prior research. In this section, we provide an overview of the composition of our data set and our methodology. We first introduce the initial sample selection of ETFs, and then outline the steps taken to process and refine the data. Lastly, we describe the analytical techniques applied in our empirical analysis.

4.1 ETF Description

The data sample used in this thesis contains ETFs that have been listed and traded on any European stock exchange between 2000-01-01 and 2024-12-31. Moreover, the data includes all ETFs that have been listed, merged, and delisted during this time period.

In order to properly investigate liquidity's effect on ETF returns and tracking errors, actively managed funds are excluded from the sample. This is due to the fact that actively managed funds are designed to create excess returns, and they typically deviate more from underlying indices, which can affect tracking errors, depending on the investment style of the ETF manager (Bae and Kim 2020). Being able to separate the liquidity effect of an ETF from the investor management style is difficult, and thus it is reasonable to exclude actively managed ETFs when doing an analysis of the liquidity effect on tracking errors. Therefore, the sample for the thesis will only include passive index-based ETFs. As shown in Figures 1 and 2, ETF listings have surged during the selected timeframe, especially since 2015, which is predominantly by funds in Luxemburg and equity ETFs.

Figure 1
ETF Counts by Domicile

Figure 1 *illustrates the total number of ETFs created and delisted over the sample period. The black line represents the development for the total number of active European ETFs each year. The coloured bands represent the contribution each domicile has to that total.*

Active European ETFs by Domicile (2000-2024)

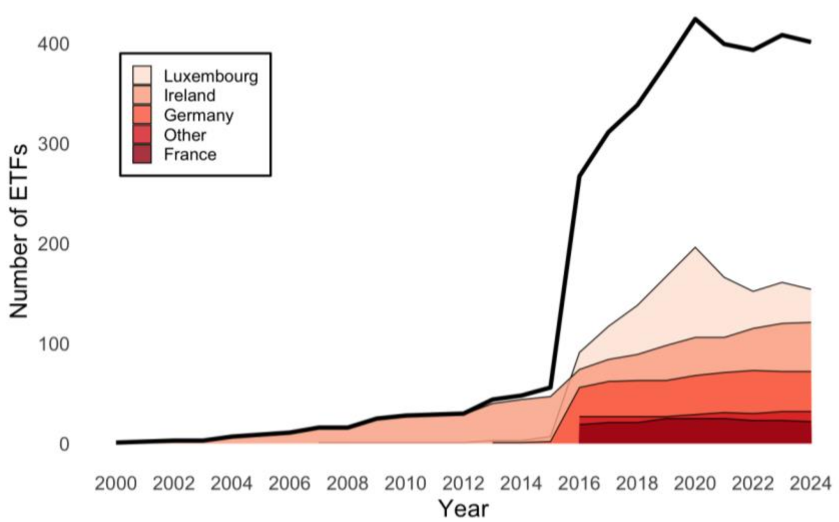
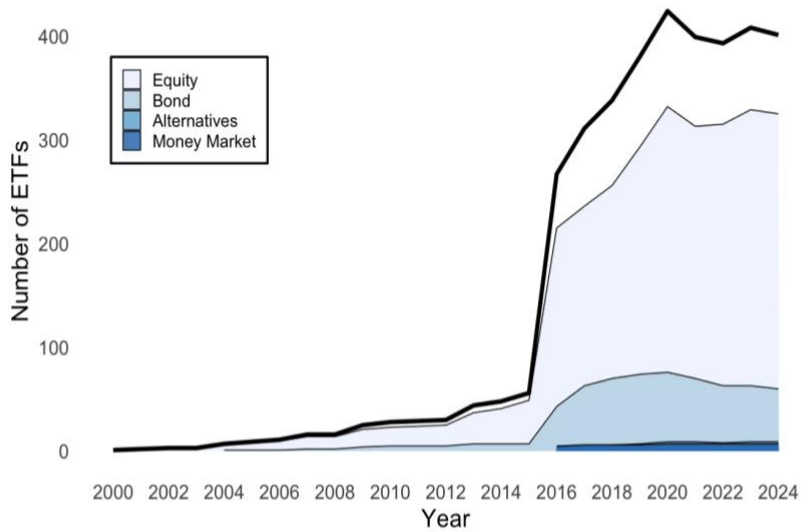


Figure 2

ETF Counts by Asset Type

Figure 2 *illustrates the total number of ETFs created and delisted over the sample period. The black line represents the development for the total number of active European ETFs each year. The coloured bands represent the contribution each ETF type has to that total.*

Active European ETFs by Asset Type (2000-2024)



4.2 Data Collection and Processing

The data on European ETFs was extracted from the Stockholm School of Economics' Refinitiv Eikon database, using a Fund screen. In the initial screen, all ETFs with a passive management approach, that have either been active, delisted or merged within Europe during the specified time period were extracted. All European countries, as well as other relevant areas such as the Eurozone were added as a geographical focus criteria in the screen. The initial sample consisted of 1330 ETFs, and included information on ETF name, ISIN code, geographical focus, asset currency code, asset type, and the name of the fund manager benchmark, also called the underlying index.

The initial sorting made after the extraction of the data was by excluding all ETFs which did not have any ISIN Code available. Without the ISIN Code, it was not possible to use formulas to extract the relevant data. It was not achievable to manually search for the ISIN Code, as each ETF might trade on several exchanges and therefore, we would perhaps not obtain ETF data in the same currency as the underlying benchmark, depending on which exchange we were to choose in the manual extraction. This reduced the sample size to 1247 ETFs.

The second sorting made was by excluding all ETFs where the underlying benchmark was not provided or available on the Lipper database, consequently reducing the sample to 847 ETFs. The reason for this is because the study aims to investigate passive ETFs, which are tracking an underlying index.

The Fscreen in Eikon is normally used to extract data on Funds, and not necessarily indices, so benchmark tickers were not always readily available. We therefore had to complement the formula-based extractions with manual retrival of the remaining underlying index tickers. This was done by searching the name of each index in the Refinitiv Eikon database and retrieving the correct ticker. For some underlying indices, neither using a formula, nor manually searching yielded a valid ticker, and the number of ETFs were thus reduced to 772.

Furthermore, to account for the fact that some ETFs might be traded in another currency than that of the underlying index, we used the underlying index ticker and a formula in excel to retrieve the currency of the benchmark index. There were 93 underlying indices which did not provide any currencies, and the sample size was hence reduced to 679 ETFs. Additionally, we computed an IF function in Excel, which would only return the ETFs where the underlying index had the same currency as the ETF itself, reducing the sample to 533 ETFs. Lastly, domiciles outside of Europe were removed to concentrate the study on the European ETF market, yielding a final sample of 494 funds.

To extract all relevant data, formulas were created through the Refinitiv Eikon add-in in Excel. Using the ISIN Code of each ETF, as well as the ticker ID on the underlying indices, we obtained daily data on the ETFs' bid price, ask price, close price, NAV and shares outstanding, as well as close price data on the ETFs' underlying indices, for the sample period 2000-01-01 to 2024-12-31. The next step was matching the trading dates in the sample with the data available, which was done through the Excel function index-match. Table 1 present a summary of the data collection process and the final sample size.

Table 1

Data Processing and Sample Size

***Table 1** illustrates the steps in processing our data, from the initial screen to the final data sample. Due to the long sample period, and many criterias needed to be fulfilled, the final sample ended up consisting of 494 European ETFs.*

Segmentation	Number of ETFs removed	Number of ETFs remaining
Initial Screen	0	1330
ISIN Code not available	83	1247
Index not available on Lipper database	400	847
No benchmark ticker available	75	772
No currency for benchmark ticker	93	679
Currency match, ETF & Index	146	533
Domicile outside Europe	39	494
Final Sample		494

4.3 Explanation of Selected Time Frame

The selected time frame covers the period 1 January 2000 until 31 December 2024 for a number of reasons. In 2000, the first European ETF was launched and started trading on Deutsche Börse (Stoxx, 2025). Hence, the sample begins in early 2000 and aligns with the introduction of the first European ETF and the start of the European ETF market.

Extending the sample to the end of 2024 ensures capturing the most recent full-year data, along with several financial disruptions, such as the COVID-19 pandemic, where the sample experienced a record in delisting rate of 9.5%. We deliberately omit the Global Financial Crisis (2007-2009) as an isolated event from our regressions, because too few ETFs were active within this time frame.

The European ETF market saw substantial growth in 2016, partially due to the introduction of new product launches, particularly smart-beta ETFs (EY, 2016). Thus, as the European ETF market took off in 2016, the COVID-19 pandemic is a better timeframe to analyze. Since the introduction of the first European ETF, the market has seen substantial growth. Table 2 provides a summary of the European ETF market development for the selected time frame.

Table 2
ETF Counts by Year

Table 2 *shows the number of ETFs that have been Created and Delisted over the studied period, based on the ETFs that have data for shares outstanding available. The column Active indicate the total number of active ETFs during that particular year, and the delisting rate reveal the percentage of delisting over the past year and is calculated by dividing the number of delisted ETFs for the current year with the number of active ETFs in the end of the past year.*

Year	Created	Delisted	Active (N)	Delisting Rate (%)
2000	1	0	1	NA
2001	1	0	2	0
2002	1	0	3	0
2003	0	0	3	0
2004	4	0	7	0
2005	2	0	9	0
2006	2	0	11	0
2007	5	0	16	0
2008	0	0	16	0
2009	9	0	25	0
2010	3	0	28	0
2011	1	0	29	0
2012	1	0	30	0
2013	14	0	44	0
2014	4	0	48	0
2015	8	0	56	0
2016	211	0	267	0
2017	44	0	311	0
2018	27	0	338	0
2019	42	1	379	0.3
2020	45	36	388	9.5
2021	11	18	381	4.6
2022	12	5	388	1.3
2023	20	15	393	3.9
2024	8	23	378	5.9

4.4 Comparison of Our Data Sample with Related Studies

Our data sample differs from that of other studies in several ways. Firstly, it covers a time-period longer than that of any other literature – 25 years compared 20 years in Bae and Kim (2020). The sample period also covers events that have not been covered by previous literature, such as the COVID-19 pandemic, as well as the war in Ukraine and the Gaza war.

Furthermore, the geographical focus in this study differs from previous research, as we only consider European ETFs. Although Tsalikis and Papadopoulous (2019) examines some European ETFs, their sample size is only 15 ETFs – 10 of whom are European.

Lastly, our final data sample includes 494 ETFs, which exceeds both the average, and the median sample size used in prior studies. Table 3 below provides a summary of the main features of the samples in related studies.

Table 3
Sample Comparison with Prior Studies

Table 3 *compares our sample to that of other relevant studies. Our sample includes more recent data, covering years well after 2012, and our sample size is both larger than the average and the median of the sample sizes for the previous studies.*

Research	Geographical Area	Time Period	Total Number of ETFs
Ackert & Tian (2008)	U.S.	2002 to 2005	28
Tsalikis & Papadopoulos (2019)	U.S & Europe	2010 to 2018	15
Bae & Kim (2020)	U.S.	1993 to 2012	1307
Our Sample	Europe	2000 to 2024	494

4.5 Measuring Illiquidity

The daily individual ETF illiquidity is proxied by the relative quoted half-spread, which is a widely accepted measure of market illiquidity and transaction costs. As defined in Equation 1, the relative quoted half-spread is calculated as half the difference between the quoted ask and bid prices, divided by the midpoint of the two quotes. The midpoint price, shown in Equation 2, is defined as the average of the bid and ask prices.

$$RQHS_t = \frac{{P^A}_t - {P^B}_t}{2\ x\ \hat{P}_t} \tag{1}$$

Where

$$\hat{P}_t = \frac{({P^A}_t + {P^B}_t)}{2} \tag{2}$$

While Bae and Kim (2020) primarily use the relative effective half-spread as their liquidity measure, the relative quoted half-spread has also been employed in their study to assess the effects of liquidity on ETF tracking errors. The relative effective half-spread requires high-frequency trade-level data, including individual trade prices and the prevailing quote midpoint at the time of each trade. This data is not available through the Refinitiv Eikon database. In contrast, the relative quoted half-spreads rely on end-of-day bid and ask prices, which are available in the Refinitiv Eikon database - making it a practical proxy for liquidity in studies based on daily data.

It is preferred to use daily level measures over monthly level measures when examining the effect of liquidity on tracking errors, as APs conduct arbitrage activities on a daily basis. These activities influence ETF pricing and can lead to daily fluctuations in tracking error. Hence, a monthly-based liquidity measure would be less suitable for capturing these short-term fluctuations in the study. (Bae and Kim, 2020)

4.6 Estimating Tracking Errors

Tracking errors are estimated using several methods. First, we estimate the return differences between (i) the ETF and its NAV, (ii) the NAV and the underlying index and (iii) the ETF and its underlying index. This is done using the daily closing price of the ETFs and the underlying indices, along with the daily NAVs. Daily tracking errors are then calculated by taking the absolute value of the return differences, as seen in Equations 3 (i), 4 (ii) and 5 (iii).

$$|ETF - NAV|_{i,t} = \left|R_{i,t}{}^{ETF} - R_{i,t}{}^{NAV}\right| \tag{3}$$

$$|NAV - IND|_{i,t} = \left|R_{i,t}{}^{NAV} - R_{i,t}{}^{IND}\right| \tag{4}$$

$$|ETF - IND|_{i,t} = \left|R_{i,t}{}^{ETF} - R_{i,t}{}^{IND}\right| \tag{5}$$

In addition, annual tracking errors are further computed using two different methods. Firstly, regression-based tracking errors are defined as the absolute difference between one and the estimated regression coefficient ($\hat{\beta}$) from regressing two return types on another. This approach is applied to the three return pairs: ETF returns versus NAV returns (ETF-NAV), (Equation 7), NAV returns versus Index returns (NAV-IND), (Equation 9), as well as ETF returns versus Index returns (ETF-IND), (Equation 11).

Equation 6 shows the ETF versus NAV tracking error, which estimates how closely the ETF market price tracks its NAV. Deviations may indicate pricing inefficiencies, market frictions or inability for the APs to keep the ETF price aligned with its NAV.

$$\theta(ETF - NAV)_{i,t} = \left|1 - \hat{\beta}_{ETF_{i,t},NAV_{i,t}}\right| \tag{6}$$

Where

$$R_{i,t}{}^{ETF} = \alpha + \beta R_{i,t}{}^{NAV} + \varepsilon_{i,t} \tag{7}$$

Equation 8 illustrates the NAV versus Index tracking error. This regression explains how well the ETF's underlying portfolio replicates the benchmark index and captures tracking issues due to index replication, fund expenses or portfolio construction.

$$\theta(NAV - IND)_{i,t} = \left|1 - \hat{\beta}_{NAV_{i,t},IND\,_{i,t}}\right| \tag{8}$$

Where

$$R_{i,t}{}^{NAV} = \alpha + \beta R\,_{i,t}{}^{IND} + \varepsilon_{i,t} \tag{9}$$

Furthermore, Equation 10 presents the ETF versus Index tracking error, which captures the total tracking error between the ETF market price and the benchmark index. This measure combines both sources of deviation – NAV misalignments, and index tracking imperfections.

$$\theta(ETF - IND)_{i,t} = \left|1 - \hat{\beta}_{ETF_{i,t},IND_{i,t}}\right| \tag{10}$$

Where

$$R_{i,t}{}^{ETF} = \alpha + \beta R_{i,t}{}^{IND} + \varepsilon_{i,t} \tag{11}$$

Secondly, the annual tracking errors are also calculated by taking the standard deviation of the return difference between the two return types multiplied by the square root of 252, to annualize the tracking errors. This method captures how volatile the tracking difference is, which means that the higher the standard deviation is, the more the ETF deviates from its benchmark over time. The standard deviation-based tracking errors can be seen in Equations 12 (i), 13 (ii), and 14 (iii).

$$\sigma(ETF - NAV)_{i,t} = StdDev\left(R_{i,t}{}^{ETF} - R_{i,t}{}^{NAV}\right) * \sqrt{252} \tag{12}$$

$$\sigma(NAV - IND)_{i,t} = StdDev\left(R_{i,t}{}^{NAV} - R_{i,t}{}^{IND}\right) * \sqrt{252} \tag{13}$$

$$\sigma(ETF - IND)_{i,t} = StdDev\left(R_{i,t}{}^{ETF} - R_{i,t}{}^{IND}\right) * \sqrt{252} \tag{14}$$

4.7 Effect of ETF Illiquidity on Tracking Errors

To investigate the effect of illiquidity on tracking errors, we conduct three panel regressions. The regression framework is based on the model used by Bae and Kim (2020), with adaptations made to account for data availability and ETF characteristics.

In all regressions, the dependent variable is tracking error, which has been estimated using three methods: (i) the regression-based method, (ii) the standard deviation-based method and (iii) the daily absolute return difference method. The main independent variable is the ETF illiquidity measure, which is proxied by the relative quoted half-spread. A higher value of this measure indicates lower liquidity, i.e. higher levels of illiquidity.

To control for other factors that may influence ETF pricing behavior, we include several control variables; NAV return volatility, Index return volatility, Log(AUM), and dummy variables for the ETF types. Annual NAV and index return volatility are computed as the standard deviation of daily log-return for each year, scaled by the square root of 252 to annualize (see Appendix). Log(AUM) is annualized by the natural log of daily shares outstanding and NAV, and averaging these on a yearly basis (see Appendix). The illiquidity is annualized through averaging the daily illiquidity across all trading days and including a small error term so the log value cannot be zero as a value (see Appendix).Volatility of the underlying index is to account for scenarios where large index movements may cause return deviations, forcing APs to rebalance more frequently, which consequently widens the gap between ETF and Index returns. Similarly, NAV return volatility is to control for deviations in the ETFs Net Asset Value, where higher volatility can signal more frequent portfolio turnover, resulting in increased tracking deviations. Log(AUM) is included to control for fund size. ETFs hold different fund characteristics, which consequently impacts liquidity in different ways. Thus, it is necessary to account for the ETF type when computing regressions on illiquidity, and hence we include asset-type dummy variables to control for these structural changes in liquidity. The dummy variables on ETF type capture the systematic differences between the varying product categories, which all have distinct trading and liquidity profiles, influencing tracking errors differently.

The three regression models are structured as follows. Model one is a yearly panel regression, where the regression-based tracking error is regressed on the ETF illiquidity measure. Furthermore, the model includes year fixed effects.

Model two is also a yearly panel regression, but with the standard deviation-based tracking error as the dependent variable. The yearly averages are computed in the same way as for model one, and it includes year fixed effects.

Model three is, unlike the other two models, a daily panel regression, where daily tracking errors are regressed on the daily ETF liquidity measure. The model includes daily fixed effects to control for market-wide shocks or events affecting all ETFs on a given day. Table 4 reports a summary of the different regression models.

Each regression model is estimated over the full sample period. Given that liquidity can deteriorate substantially during periods of market stress, we also analyze subsamples corresponding to major market crises, specifically the COVID-19 pandemic period between 2020 and 2021 (de Vette et al. 2023). This allows us to assess whether the impact of liquidity on tracking error becomes more pronounced during crisis periods.

In our panel data, there is a risk that errors may be correlated within each ETF over time, as well as across different ETFs within the same year due to common shocks, such as events that affect many, or all ETFs simultaneously – for instance geopolitical events or macroeconomic announcements. Such crisises can impact the returns of numerous ETFs in the same year, which introduces correlation in their residuals. By simply ignoring these time-based dependencies, there is a risk of underestimating standard errors, which in turn can inflate the t-statistics and thus increase the risk of falsely identifying variables as statistically significant. To address this issue, we use double-clustered standard errors, clustering by both fund and time frame; yearly for the two first models and daily for the third model. This method ensures that our significance levels reflect the true variability in the data and that they are not distorted by underlying dependencies in the sample data.

To assess the significance of the observed effect, we conduct a hypothesis test. In this context, we use a two-tailed test-statistic to evaluate whether the estimated coefficients from the regression models are statistically different from zero. The test-statistic is used to account for both positive and negative deviations from the null. The null hypothesis is that the beta coefficient for the relative quoted half-spread (our proxy for ETF illiquidity), is equal to zero – implying that changes in ETF liquidity have no effect on tracking error. The alternative hypothesis is that the coefficient differs from zero, and if t exceeds the critical threshold we reject the null and conclude a statistically significant relationship. Since our sample size is large (N > 30) if follows from the Central Limit Theorem (CLT) that the estimated coefficient $\hat{\beta}$ is approximately normally distributed under the null hypothesis, as seen in Equation 15. The two-tailed test-statistic is specified in Equation 16, where the standard error of the estimated coefficient is adjusted using double-clustered standard errors.

$$\hat{\beta} \sim N\left[0, Var\left(\hat{\beta}\right)\right] \tag{15}$$

$$t = \frac{\hat{\beta}}{SE\left(\hat{\beta}\right)} \tag{16}$$

Table 4
Regression Summary

***Table 4** compares the three different regressions used to investigate the effect of liquidity on tracking errors. Number one is a yearly regression, based on the regression-based tracking errors, while number two is based on the standard deviation-based tracking errors. In the third regression, daily regressions of the daily tracking errors are made on the ETF illiquidity. All regressions employ the same control variables. Instead of double clustering the standard errors on the fund and year level, like regression one and two, the third regression double clusters the standard errors on fund and day level.*

Number	1	2	3
Tracking Errors	Regression-based	Standard Deviation-based	Daily Absolute Return Difference
Regression type	Yearly panel regression	Yearly panel regression	Daily panel regression
Fixed Effects	Year	Year	Day
Standard Errors	Double Clustered: Fund & Year Level	Double Clustered: Fund & Year Level	Double Clustered: Fund & Day Level
Control Variables	NAV Return Volatility Index Return Volatility Log (AUM) Dummy variable: ETF Type[1]	NAV Return Volatility Index Return Volatility Log (AUM) Dummy variable: ETF Type[2]	NAV Return Volatility Index Return Volatility Log (AUM) Dummy variable: ETF Type[3]

4.8 Data Limitations

We identify several limitations in our sample data that may affect the robustness and generalizability of our findings.

First, although our final data sample includes 494 ETFs, not all of them contain information on all relevant variables. As a result, only the ETFs with sufficiently complete data, can be used in key parts of the analysis, such as the regression models. This means that the effective sample may be skewed, for instance toward larger or more liquid ETFs that are more likely to have a more complete data coverage. Smaller ETFs, or those with shorter lifespans or lower trading activity, may therefore be disproportionally excluded from the analytical results, even though they are present in the overall dataset.

Second, we restricted our sample to ETFs trading in the same currency as their underlying benchmark. While this enhances comparability between ETF and index results, it excludes ETFs operating in multi-currency environments. This introduces a potential sampling bias, as the excluded 146 ETFs represent approximately 11% of the initial sample of 1330 ETFs.

[1] ETF Type: Bond, Equity and Money Market. The full sample contains Alternatives, Bond, Equity and Money Market.
[2] ETF Type: Bond, Equity and Money Market. The full sample contains Alternatives, Bond, Equity and Money Market.
[3] ETF Type: Bond, Equity and Money Market. The full sample contains Alternatives, Bond, Equity and Money Market.

Third, during data collection, we relied in part on manual retrieval of underlying index tickers when automated formula-based extraction failed. This process may have introduced human error, potentially affecting the accuracy and completeness of the final sample.

Although Refinitiv Eikon is a well-established financial data provider, its coverage of European ETFs is not exhaustive. As evidenced during the data screening process, it does not cover nor have data on all European ETFs. This creates a risk of coverage bias and could limit the generalizability of the study's conclusions to the broader European ETF market.

Lastly, in the subsample analyses (particularly for the COVID-19 period), the annual regressions based on regression- and standard deviation-based tracking errors, includes only 244 ETF-year observations. Such small sample can lead to imprecise coefficient estimates, larger standard errors and higher sensitivity to outliers. While the crisis period results offer valuable insights into stressed market dynamics, they should be interpreted as exploratory rather than definitive.

5 Empirical Findings and Discussion

This section presents the main empirical results of the study. We begin with a descriptive overview of the annual tracking errors observed in the ETF sample, followed by an analysis of distributions between ETF, NAV and Index returns. Next, we examine the time series relationship between illiquidity and tracking deviations, and finally, we assess the influence of ETF illiquidity on tracking errors though a series of regression models.

Table 5 presents summary statistics for the two types of annual tracking errors. The standard deviation-based tracking errors between the ETF and the NAV returns (σ (ETF−NAV)) suggests that ETF prices, on average, relatively closely, but not perfectly follow their NAVs. This suggests that market pricing and arbitrage by APs generally are effective. For the NAV versus Index standard deviation measure, the average tracking errors are lower, but the standard deviation is higher, which implies that some ETFs replicate their indices very well, while others have larger tracking issues. The mean tracking error between ETFs and the Index (σ (ETF−IND)) is the highest among the standard deviation-based metrics at 0.38%. This is expected, as it reflects the total tracking error and shows that the deviations add up when both the ETF and the NAV do not track the underlying assets or index correctly.

The regression-based tracking errors (θ), are overall larger than the standard deviation-based tracking errors (σ). While the standard deviation-based tracking errors (σ) reflect the average deviation in returns, the regression-based tracking errors (θ) captures structural mismatches in returns, where a high value reflects that return differences consistently under- or overreact to changes in the underlying measure. This means that while an ETF may track its NAV or underlying index well on a day-to-day basis, and thus have a low value of standard deviation-based tracking errors, it could fail to proportionally replicate the benchmark over time, and hence have a high value of the regression-based tracking errors. This, based on the summary statistics, seems to be the case for the European ETFs.

The mean of the regression-based tracking error for the ETF versus NAV return difference is about 8.63%, indicating that ETF prices do not move perfectly in sync with their NAVs. Furthermore, the standard deviation is high, at 20.22%, meaning that while some ETFs track their NAVs rather closely, other ETFs have substantial variation. On average, NAV returns deviate by 6.88% from the benchmark index. This reflects tracking inefficiencies at the portfolio level. The high standard deviation for NAV versus Index (20.68%) signals that some ETFs replicate indices well, while others deviate significantly. Lastly, the overall tracking error between the ETF return and the Index return is, as expected, the highest value since it includes both ETF versus NAV tracking issues, and NAV versus Index replication issues. The relatively high mean and large standard deviation of θ(ETF−IND) indicate that many ETFs deviate from the indices, both on average and in terms of volatility.

For the summary statistics, these results are in line with Bae and Kim (2020). In general, ETFs track NAVs better than they track their underlying indices, whether it is measured by regression-based or standard deviation-based tracking errors. NAVs typically track the indicies more closely than ETFs, though some tracking error remains. This suggests that ETF returns are prone to deviate from their index or NAV return due to external market factors that are beyond the control of fund managers. Furthermore, the highest average tracking error and variability occur between the ETFs and the indicies, reflecting the combined effects of both ETF return deviation and the NAV replication problems.

Table 5

Summary Statistics for Annual Tracking Errors

Table 5 *presents a summary of the estimated annual tracking errors for the ETFs, with both the regression-based tracking errors, as well as the standard deviation-based tracking errors, measured from their inception date until either the end of the sample period or their delisting date.*

Tracking Error	Mean	Standard Deviation
σ (ETF−NAV)	0.32%	0.24%
σ (NAV−IND)	0.19%	0.54%
σ (ETF−IND)	0.38%	0.55%
θ (ETF−NAV)	8.63%	20.22%
θ (NAV−IND)	6.88%	20.68%
θ (ETF−IND)	13.86%	30.20%

Table 6 further displays the average cross-sectional correlations of the estimated annual tracking errors. One interesting result is the strong correlation between σ(NAV−IND) and σ(ETF−IND) of 0.915. This suggests that the standard deviation-based tracking error for NAV versus Index is a major driver of the standard-deviation based tracking error for ETF versus Index. The results also show moderate correlations between σ(ETF−NAV) and σ(ETF−IND) (0.404), which indicates that ETF price deviations from its NAVs contribute somewhat to the total tracking error (ETF-IND), but less than the NAV versus Index mismatch. For the regression-based tracking errors, the ETF versus NAV tracking errors correlate moderately with the ETF versus Index tracking errors (0.652), meaning that if an ETF does not move toghether with its NAV, it is also likely that they will not move in sync with its index. Lastly, the θ(NAV−IND) and θ(ETF−IND) have a strong correlation of 0.732, where a poor NAV versus Index replication strongly affects the ETF versus Index tracking errors. The regression-based tracking errors are rather strongly correlated with the standard deviation-based tracking errors for each measure (0.541 for ETF-NAV, 0.376 for NAV-IND and 0.304 for ETF-IND), which, similarly to Bae and Kim (2020), indicates that the two types of tracking errors used are reasonably robust and in line with each other.

Table 6

Cross-Sectional Correlations for Annual Tracking Errors

Table 6 *presents the average cross-sectional correlations of the estimated annual tracking errors for the ETFs, measured from their inception date until either the end of the sample period or their delisting date.*

	σ(ETF−NAV)	σ(NAV−IND)	σ(ETF−IND)	θ(ETF−NAV)	θ(NAV−IND)	θ(ETF−IND)
σ(ETF−NAV)	1.000					
σ(NAV−IND)	0.187	1.000				
σ(ETF−IND)	0.404	0.915	1.000			
θ(ETF−NAV)	0.541	0.036	0.191	1.000		
θ(NAV−IND)	-0.136	0.376	0.275	0.066	1.000	
θ(ETF−IND)	0.194	0.226	0.304	0.652	0.732	1.000

Although the preceding tables provide summary statistics and correlations that describe the overall tracking performance of ETFs in the sample, they do not show how tracking errors vary across different types of ETFs, or over the sample period. To further investigate these dynamics, a graphical analysis of the cross-sectional relationships between ETF returns and NAV returns (Panel A), as well as between NAV returns and Index returns (Panel B) has been employed. These plots are shown for (i) the full sample over the entire sample period, and (ii) the oldest ETF (iShares STOXX Europe 600 Health Care (DE)) in the sample. Under ideal market conditions – without frictions, APs' actions should ensure that the daily returns of ETFs, their NAVs and the underlying index align closely, since any deviations would be eliminated through arbitrage.

Figure 3 presents the cross-sectional relationships between ETF returns and NAV returns (Panel A) and between NAV returns and benchmark index returns (Panel B) for the full sample. Each black dot represents the time series average of daily returns for an individual ETF, between its inception date until its delisting date, or the end of the sample period. The red line shows the fitted regression line, while the green dotted line represents the 45-degree line, which serves as a benchmark for perfect tracking – indicating equality between the two return-series.

In Panel A, the fitted slope coefficient is 0.85, which means that the ETF returns reflect approximately 85% of the variation in NAV returns. This suggests that ETF market prices do not completely incorporate changes in their NAVs, which means that there is some degree of market pricing inefficiency in the sample. The R^2 of 60.6% implies that NAV returns explain around 61% of the cross-sectional variation in ETF returns.

In Panel B, the fitted coefficient is lower at 0.70. This means that NAV returns reflect only 70% of changes in index returns, and thus there are larger tracking inefficiencies at the portfolio replication level (NAV versus Index), compared to the market pricing level (ETF versus NAV). This notion aligns with the strong correlation (0.915) between σ(NAV−IND) and σ(ETF−IND) from Table 6. The R^2 is higher at 66.5%, suggesting that index returns explain more of the variation in NAV returns than NAV returns explain ETF returns.

Figure 3
Return Distributions Across ETFs, NAVs and Indices

Figure 3 *illustrates the cross-sectional relationships between ETF returns and NAV returns (Panel A), and between NAV returns and benchmark index returns (Panel B), for all ETFs in the sample. Each black dot represents the time series average of daily returns for an individual ETF between its inception date until its delisting date or the end of the sample period. The red line shows the fitted regression line, while the green dotted line represents the 45-degree line, which serves as a benchmark for perfect tracking.*

Panel A: ETF vs. NAV **Panel B: NAV vs. Index**

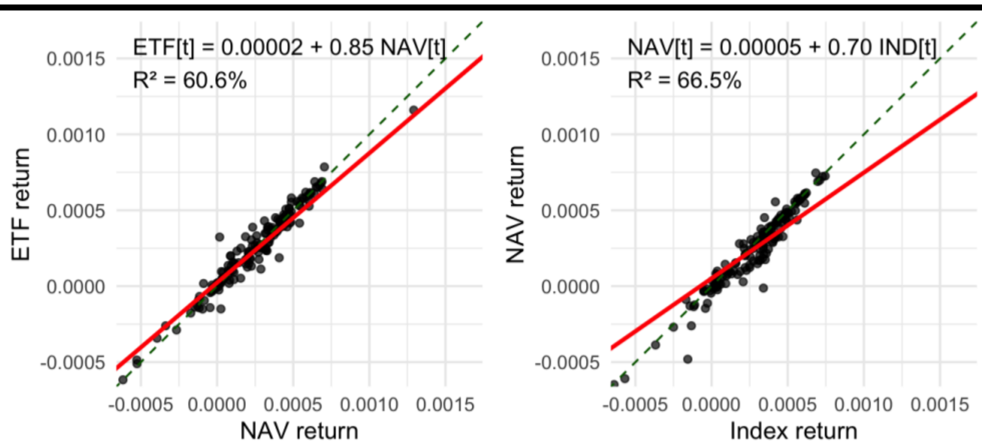


Figure 4 shows the same cross-sectional panels as Figure 3, but instead of showing the entire sample, it only presents the returns for an individual ETF – namely the oldest one in the sample. The oldest ETF in the sample is *iShares STOXX Europe 600 Health Care (DE),* which was incepted on the 25th of April 2001. Overall, the ETF show very strong tracking performance at both the market pricing level (ETF versus NAV), as well as the portfolio replication level (NAV versus Index). In Panel A, the fitted coefficient is 0.95. This indicates that the ETF returns respond to 95% of the variation in NAV returns, which is a high degree of co-movement. Furthermore, the R^2 is 86.5%, which indicates that the ETF pricing closely reflects that of the underlying assets. This suggests that arbitrage mechanisms function effectively and that *iShares STOXX Europe 600 Health Care (DE)* trades with relatively low pricing deviations – likely because of its long history and familiarity in the market.

Panel B shows that nearly all index return movements are reflected in the NAV, as seen in the high fitted coefficient of 0.96. Additionally, the high R^2 (92.5%) indicates that the ETF's portfolio replicates the underlying index closely, with only small deviations over time.

The results demonstrate significant variation in ETF tracking performance across the full sample. While the oldest ETF exhibit fairly high alignment with both its

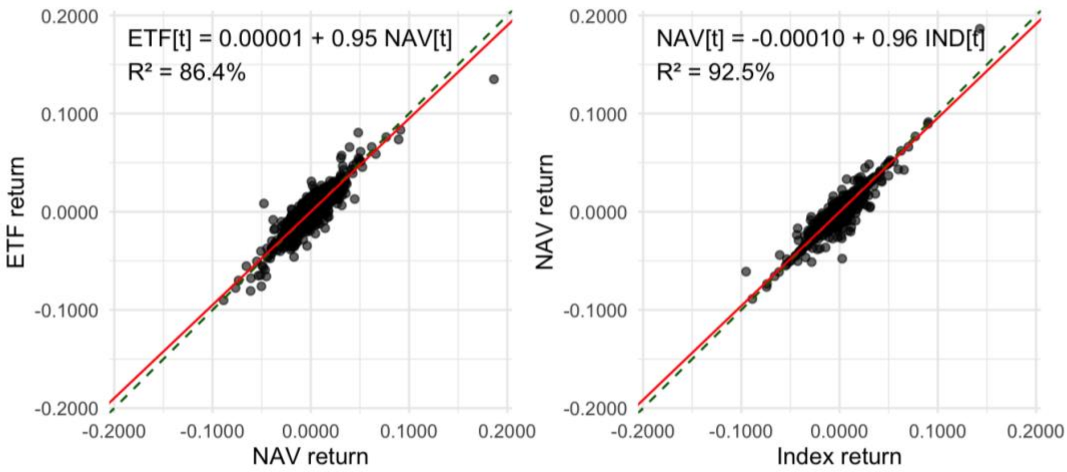
NAV and index returns, the full sample has notable deviations, especially at the portfolio replication level (NAV versus Index). These results are in line with those of Bae and Kim (2020), although the European ETFs in our sample generally exhibit weaker tracking quality than the U.S. ETFs in their study.

Figure 4

Return Distributions Across ETF, NAV and Index for the Oldest ETF

Figure 4 *illustrates cross-sectional relationships between ETF returns and NAV returns (Panel A) and between NAV returns and benchmark index returns (Panel B) for the ETF "iShares STOXX Europe 600 Health Care (DE)", which tracks the underlying index "STOXX Europe 600 Health Care TR EUR". The red line depicts the fitted regression line, and the black dots represents the average of daily returns from 25/04/2001 to the end of the sample period. The green dotted line is the 45-degree line, which serves as a benchmark for perfect tracking.*

Panel A: ETF vs. NAV	**Panel B: NAV vs. Index**



While the prior analysis focused on cross-sectional variation in tracking performance, it does not account for market frictions. To capture these effects, the time series relationship between ETF illiquidity and tracking deviations is examined. Figure 5 illustrates this connection by plotting historical time series of illiquidity and the absolute values of return differences from 2008 to 2024. The graph shows that both ETF-NAV tracking deviations and illiquidity increase substantially during periods of financial stress. During the COVID-19 pandemic in early 2020, there is a sharp and simultaneous spike in both the ETF-NAV gap and the illiquidity measure, which signals that there are substantial pricing frictions and weakened arbitrage.

This pattern is consistent with the concept of *commonality in liquidity* described by Holden et. al (2014), where liquidity conditions co-move across assets. During periods of market volatility, declines in both ETF and underlying asset liquidity presumably reinforce one another, which leads to increased pricing difficulties and tracking errors.

In addition to the spike during the COVID-19 pandemic, similar patterns are visible during other periods of market crisis, for instance during the 2008-2009 Global

Financial Crisis and the 2009-2010 European Sovereign Debt Crisis. Each of these periods are marked by a simultaneous increase in the ETF-NAV gap and market illiquidity. These results thus suggest that under extreme market conditions, ETF prices are more likely to diverge from their underlying NAVs because of reduced liquidity and wider trading costs. Although the NAV-Index gaps generally are smaller and more stable, they also exhibit a temporary increase during these periods, which reflects disruptions in portfolio replication.

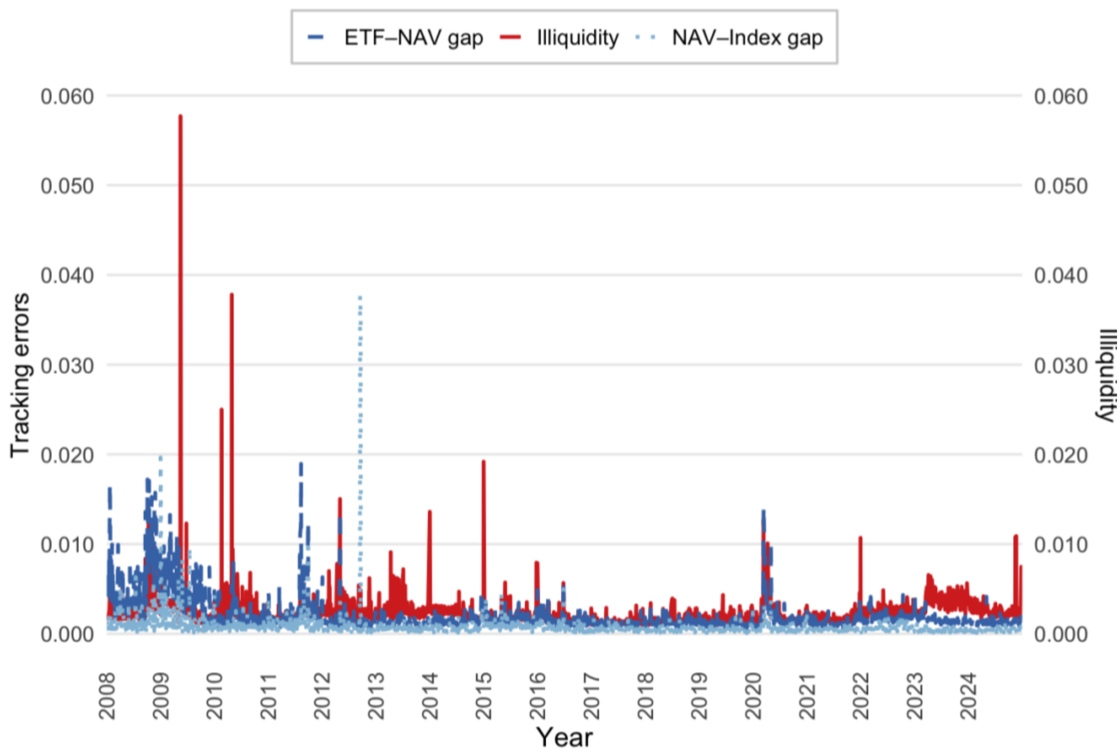
Over the longer term, both illiquidity and tracking deviations show a downward trend after 2012. This indicates market maturation and could be a result of improving liquidity conditions, or more efficient ETF pricing in European markets. It could also be a result of a larger number of ETFs being available in the market, which flattens out averages.

Figure 5

Daily Time Series of Illiquidity Measures and Return Deviations

Figure 5 *depicts the connection between ETF illiquidity and ETF tracking errors over time, where historical time series of illiquidity and absolute values of return differences from 2008 to 2024 are plotted. The red line represents the average illiquidity, which is the cross-sectional average of the relative quoted half-spread of all ETFs. Furthermore, the blue line represents the average absolute value of return differences between ETF and NAV returns (ETF-NAV). Lastly, the light blue dotted line is the average absolute value of return differences between NAV and Index returns (NAV-IND).*

Illiquidity and Return Gaps (2008-2024)



The daily time series of illiquidity measures and return deviations provides visuals that points towards a positive relationship between ETF illiquidity and tracking errors, especially during periods of market stress. However, this visual analysis does not control for other potentially confounding factors. To formally examine this relationship, panel regressions of the regression-based tracking errors on ETF illiquidity are estimated.

Table 7 summarises the results for both the full sample, as well as the COVID-19 subperiod, and allows for assessment of whether illiquidity has a statistically significant impact on regression-based ETF tracking errors, and if this relationship is intensified during a specific period with systemic volatility (the COVID-19 subperiod). The table presents some interesting results. The beta-coefficient for the Log ETF illiquidity is positive and statistically significant only for θ(ETF–NAV) (ß = 0.040, t = 4.215), and non-significant for the NAV versus Index (θ(NAV–IND), (ß = 0.004, t = 0.159)), and ETF versus Index tracking errors (θ(ETF–IND), (ß = 0.042, t = 1.378)). This signals that the illiquidity is primarily associated with ETF pricing inefficiencies between the ETF prices and their NAVs, and not with their replication qualities of NAVs on indices. In conditions with more liquidity, the arbitrage mechanisms are thus more effective in keeping ETF prices aligned with NAV than they are in replicating the underlying index. This implies that market frictions primarily impact ETF pricing, and not portfolio replication (NAV versus Index). The θ(NAV–IND) and θ(ETF–IND) are in this regression not significantly affected by the illiquidity measure, which suggests that replication issues of the NAV on the index are partly independent of trading frictions.

Another interesting result is that the money market ETFs have strongly positive and significant coefficients for θ(ETF–NAV) (ß = 0.223, t = 3.359) and θ(ETF–IND) (ß = 0.352, t = 2.800). This means that money market ETFs consistently show higher pricing and total tracking errors. This is likely due to infrequent trading in the underlying instruments.

In the COVID-19 subsample, the coefficients for NAV return volatility and Index return volatility are exceptionally large in magnitude. As an example, the coefficient on NAV return volatility is ß = 166.873 for θ(NAV–IND) and ß = 188.559 for θ(ETF–IND). Additionally, the Index return volatility has similarly large and opposite-signed coefficients; ß = −167.104 for θ(NAV–IND), and ß = −189.346 for θ(ETF–IND). Albeit these large values, the test statistics associated with these coefficients do not exceed the standard significance thresholds. These results thus raise concerns about multicollinearity, especially given the smaller number of observations in the COVID-19 subsample (244 observations). Supporting this, the cross-sectional correlation matrix in Table 6 reveals a very high correlation (0.915) between σ(NAV−IND) and σ(ETF−IND), and strong correlations between θ(NAV−IND) and θ(ETF−IND) (0.732), as well as between θ(ETF−NAV) and θ(ETF−IND) (0.652). These strong correlations indicate that the components of tracking error are not fully independent, which makes it difficult is isolate the effects of individual volatility variables (NAV return volatilty and Index return volatility) without introducing multicollinearity. As a result, the large coefficients observed during the COVID-19 period are likely attributable to shared underlying variation rather than independent expanatory effect in the regression model. We thus interpret these results with caution.

Table 7

ETF Illiquidity and Regression-based Tracking Errors

***Table** 7 presents coefficient estimates for annual regressions of ETF tracking errors on ETF illiquidity. The dependent variable is the regression-based tracking error, which is calculated by taking the absolute difference between one and the estimated regression coefficient ($\hat{\beta}$) from regressing two return types on another. This approach is applied to the three return pairs: ETF returns versus NAV returns (Columns 1 and 4), NAV returns versus Index returns (Columns 2 and 5), as well as ETF returns versus Index returns (Columns 3 and 6). The estimations for the full sample can be seen in Columns 1 to 3, while Columns 4 to 6 show the estimation results for the COVID-19 period. The most important independent variable is the Log ETF Illiquidity, which is computed by taking the logaritm of the relative quoted half-spread, plus an error term of 10^{-4} to ensure that log(value) never is zero. Additional variables are defined in the Appendix B. All regressions include year fixed effects, and test statistics based on double clustered standard errors at the fund and year level are shown in parentheses. * denote significance at the 10% level, ** denote significance at the 5% level, and *** denote significance at the 1% level.*

	All Sample			**COVID-19**		
	θ(ETF–NAV)	**θ(NAV–IND)**	**θ(ETF–IND)**	**θ(ETF–NAV)**	**θ(NAV–IND)**	**θ(ETF–IND)**
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Log ETF Illiquidity	0.040***	0.004	0.042	0.035	0.076	0.106
	(4.215)	(0.159)	(1.378)	(1.680)	(0.517)	(0.560)
NAV Return Volatility	0.805	19.069	21.032	−1.403**	166.873	188.559
	(0.975)	(1.064)	(1.064)	(−29.580)	(1.634)	(1.649)
Index Return Volatility	−1.025	−20.012	−22.711	1.265**	−167.104	−189.346
	(−1.269)	(−1.084)	(−1.106)	(22.719)	(−1.634)	(−1.653)
Log(AUM)	−0.002	−0.051	−0.052	−0.000	0.076	0.090
	(−0.456)	(−0.997)	(−0.893)	(−0.077)	(2.078)	(1.520)
ETF Type: Bond	0.055*	0.442	0.470	0.084	0.313	0.357
	(2.042)	(1.386)	(1.363)	(1.278)	(1.394)	(1.530)
ETF Type: Equity	−0.023	0.126	0.125	−0.075	−0.145	−0.203
	(−1.221)	(1.158)	(1.049)	(−3.141)	(−0.438)	(−0.564)
ETF Type: Money Market	0.223***	0.130	0.352**	0.419	−0.007	0.616
	(3.359)	(1.429)	(2.800)	(3.880)	(−0.043)	(5.804)
Number of Observations	1069	1069	1069	244	244	244
R^2	0.198	0.063	0.075	0.263	0.271	0.264
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

As a robustness check, the panel regression model for the regression-based tracking errors is re-estimated using only log ETF illiquidity as the independent variable. All other variables, such as log AUM, Index return volatility, NAV return volatility and ETF type dummies have been omitted from the model. The simplified model allows us to assess whether ETF illiquidity alone explains variation in tracking errors, and whether previous results have been driven by confounding factors or multicollinearity. The results, reported in Table 8 show that the log ETF illiquidity is not significantly associated with θ(ETF–NAV) or θ(ETF–IND), neither in the full sample, nor in the COVID-19 subsample. For the θ(NAV–IND) tracking error, there is a negative and weakly significant coefficient for the full sample (ß = −0.077, t = −1.809). This value is rather small in magnitude, and the statistical significance of it disappears during the COVID-19 period.

Of particular note, the R^2 values are considerably smaller in the simplified model; ranging from 0.005 to 0.030, compared to the specified model where the R^2 values range from 0.063 to 0.271. This means that explanatory power of the independent variable log ETF Illiquidity is somewhat stronger for the specified model with control variables, compared to this regression. Thus, this points toward the inclusion of controls help isolate the effect of ETF illiquidity more clearly.

Table 8

ETF Illiquidity and Regression-based Tracking Errors – Diagnostic Test

Table 8 *presents coefficient estimates for annual regressions of ETF tracking errors on ETF illiquidity. The dependent variable is the regression-based tracking error, which is calculated by taking the absolute difference between one and the estimated regression coefficient ($\hat{\beta}$) from regressing two return types on another. This approach is applied to the three return pairs: ETF returns versus NAV returns (Columns 1 and 4), NAV returns versus Index returns (Columns 2 and 5), as well as ETF returns versus Index returns (Columns 3 and 6). The estimations for the full sample can be seen in Columns 1 to 3, while Columns 4 to 6 show the estimation results for the COVID-19 period. The independent variable is the Log ETF Illiquidity, which is computed by taking the logaritm of the relative quoted half-spread, plus an error term of 10^{-4} to ensure that log(value) never is zero. All regressions include year fixed effects, and test statistics based on double clustered standard errors at the fund and year level are shown in parentheses. * denote significance at the 10% level, ** denote significance at the 5% level, and *** denote significance at the 1% level.*

	All Sample			**COVID-19**		
	θ(ETF–NAV)	**θ(NAV–IND)**	**θ(ETF–IND)**	**θ(ETF–NAV)**	**θ(NAV–IND)**	**θ(ETF–IND)**
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Log ETF Illiquidity	−0.002	−0.077*	−0.079	−0.043	−0.183	−0.277
	(−0.127)	(−1.809)	(−1.546)	(−1.674)	(−0.928)	(−1.031)
Number of Observations	1535	1535	1535	255	255	255
R^2	0.030	0.009	0.010	0.026	0.005	0.008
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

In the second regression model, three types of standard deviation-based tracking errors; ETF versus NAV, NAV versus Index, and ETF versus Index, are regressed on the log ETF illiquidity measure and other control variables for both the full sample and the COVID-19 period. Table 9 summarises the results. There are several interesting points to be made. The first key finding is that the coefficient on log ETF illiquidity is positive and highly significant for the standard deviation-based tracking errors on ETF versus NAV returns (σ(ETF–NAV), (ß = 0.021, t = 4.804)) and for the ETF versus Index returns (σ(ETF–IND), (ß = 0.020, t = 4.691)). For the NAV versus Index tracking error, the coefficient is still positive (ß = 0.003) although not significant (t = 1.201). These findings are in line with those of Bae and Kim (2020) and suggests that illiquidity has a strong effect on tracking errors where ETF returns are involved and implies that ETF versus NAV returns and ETF versus Index returns are sensitive to trading frictions, such as transaction fees or higher bid-ask spreads.

The table also shows that NAV and Index return volatility are strongly associated with the NAV versus Index tracking error. For the full sample, the NAV return volatility is positive and significant for σ(NAV–IND) (ß = 1.479, t = 5.491) and the Index return volatility is negative and significant for σ(NAV–IND) (ß = -1.385, t = -5.092). The positive coefficient for the NAV return volatility could be expected, as more volatile NAV returns will affect the NAV's deviation from the underlying index and result in higher tracking errors.

Furthermore, the negative sign on the coefficient for Index return volatility signals that as index return volatility increases, the standard deviation-based tracking error between NAV and index decreases. This may seem counterintuitive, since we would expect tracking errors to increase in volatile markets. One possible explanation to this result is that when index returns are less volatile, the NAVs might not exactly follow the index due to pricing practices and thus the tracking error increases. In contrast, when index returns are more volatile, NAVs might be forced to respond more directly, and thus, the return gap decreases, leading to lower tracking errors.

Lastly, the bond type and money market type ETFs exhibit significantly higher tracking errors. For the full sample, the coefficients for the Bond ETFs are positive and moderately significant for the σ(ETF–NAV) and the σ(ETF–IND) tracking errors; (ß = 0.011, t = 2.406) for σ(ETF–NAV) and (ß = 0.012, t = 2.149) for σ(ETF–IND). During the COVID-19 period, the coefficients are positive and larger in magnitude, although not significant using the standard thresholds for statistical significance. The coefficients for the money market type ETFs are positive and somewhat significant for the σ(ETF–NAV) and the σ(ETF–IND) tracking errors; (ß = 0.019, t = 2.548) for σ(ETF–NAV) and (ß = 0.017, t = 2.046) for σ(ETF–IND). These results are in line with the findings in Table 7 and could be due to bond and money market ETFs often investing in underlying instruments that trade less frequently or are harder to price in real time. As a result, these ETFs may experience wider bid-ask spreads and limited arbitrage opportunities, both of which contribute to elevated tracking errors.

Although the COVID-19 subsample regression provides insight into tracking behavior during a period of heightened market stress, the number of observations (244) in the subsample are considerably smaller than in the full sample (1069). As a result, the threshold for statistical significance is higher due to the reduced degrees of freedom, which means that larger test statistics are needed to meet the ordinary significance levels. Even though some coefficients in the COVID-19 regressions may seem meaningful, they do not reach statistical significance at the standard thresholds and are thus not the focus of interpretation.

Table 9

ETF Illiquidity and Standard Deviation-based Tracking Errors

***Table 9** presents coefficient estimates for annual regressions of ETF tracking errors on ETF illiquidity. The dependent variable is the standard deviation-based tracking error, which is calculated by taking the standard deviation of the return difference between the two return types. This approach is applied to the three return pairs: ETF returns versus NAV returns (Columns 1 and 4), NAV returns versus Index returns (Columns 2 and 5), as well as ETF returns versus Index returns (Columns 3 and 6). The estimations for the full sample can be seen in Columns 1 to 3, while Columns 4 to 6 show the estimation results for the COVID-19 period. The most important independent variable is the Log ETF Illiquidity, which is computed by taking the logaritm of the relative quoted half-spread, plus an error term of 10^{-4} to ensure that log(value) never is zero. Additional variables are defined in the Appendix B. All regressions include year fixed effects, and test statistics based on double clustered standard errors at the fund and year level are shown in parentheses. * denote significance at the 10% level, ** denote significance at the 5% level, and *** denote significance at the 1% level.*

	All Sample			**COVID-19**		
Variable	**σ(ETF–NAV)** (1)	**σ(NAV–IND)** (2)	**σ(ETF–IND)** (3)	**σ(ETF–NAV)** (4)	**σ(NAV–IND)** (5)	**σ(ETF–IND)** (6)
Log ETF Illiquidity	0.021***	0.003	0.020***	0.024	0.001	0.023
	(4.804)	(1.201)	(4.691)	(5.958)	(0.180)	(5.210)
NAV Return Volatility	0.748	1.479***	0.602	−0.210	0.233	0.143
	(1.628)	(5.491)	(1.610)	(−0.728)	(0.254)	(0.161)
Index Return Volatility	−0.638	−1.385***	−0.508	0.319	−0.206	−0.028
	(−1.429)	(−5.092)	(−1.388)	(1.297)	(−0.227)	(−0.032)
Log(AUM)	0.001	0.005**	0.003	−0.001	0.003	0.001
	(0.775)	(2.448)	(1.527)	(−0.360)	(1.532)	(0.596)
ETF Type: Bond	0.011**	0.003	0.012**	0.017	0.005	0.017
	(2.406)	(0.651)	(2.149)	(2.322)	(0.767)	(1.986)
ETF Type: Equity	−0.003	−0.000	0.003	−0.004	0.009	0.000
	(−0.474)	(−0.046)	(0.424)	(−0.385)	(1.271)	(0.007)
ETF Type: Money Market	0.019**	0.013	0.017*	0.017	−0.001	0.015
	(2.548)	(1.523)	(2.046)	(2.078)	(−0.207)	(2.020)
Number of Observations	1069	1069	1069	244	244	244
R^2	0.445	0.443	0.375	0.404	0.075	0.362
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Table 10 presents the diagnostic results for the standard deviation-based tracking error model, using log ETF illiquidity as the sole explanatory variable. The model confirms that the relationship between ETF illiquidity and ETF versus NAV tracking errors is robust, since the effect remains statistically significant even in the absence of controls. This suggests that illiquidity levels are a key driver of pricing deviations between ETFs and their NAVs.

However, the fully specified model offers a more complete picture by accounting for other structural and market-related influences, such as the fund size and the ETF type. This is reflected in the higher R^2 values, ranging from 0.075 to 0.445, compared to the simplified model with a range of 0.020 to 0.441. Examining for instance the ETF versus Index tracking error for the entire sample period, we see that the explanatory power increases from 0.065 to 0.375 in the fully specified model. In summary, the consistency across the two models thus supports the reliability of the main result, while the full model provides higher explanatory power.

Table 10

ETF Illiquidity and Standard Deviation-based Tracking Errors – Diagnostic Test

Table 10 *presents coefficient estimates for annual regressions of ETF tracking errors on ETF illiquidity. The dependent variable is the standard deviation-based tracking error, which is calculated by taking the standard deviation of the return difference between the two return types. This approach is applied to the three return pairs: ETF returns versus NAV returns (Columns 1 and 4), NAV returns versus Index returns (Columns 2 and 5), as well as ETF returns versus Index returns (Columns 3 and 6). The estimations for the full sample can be seen in Columns 1 to 3, while Columns 4 to 6 show the estimation results for the COVID-19 period. The independent variable is the Log ETF Illiquidity, which is computed by taking the logaritm of the relative quoted half-spread, plus an error term of* 10^{-4} *to ensure that log(value) never is zero. All regressions include year fixed effects, and test statistics based on double clustered standard errors at the fund and year level are shown in parentheses. * denote significance at the 10% level, ** denote significance at the 5% level, and *** denote significance at the 1% level.*

	All Sample			**COVID-19**		
	σ(ETF–NAV)	**σ(NAV–IND)**	**σ(ETF–IND)**	**σ(NAV–IND)**	**σ(ETF–IND)**	**σ(ETF–NAV)**
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Log ETF Illiquidity	0.023***	−0.004	0.014	0.004	0.027	0.027
	(6.332)	(−0.456)	(1.658)	(2.060)	(6.002)	(5.677)
Number of Observations	1535	1535	1535	255	255	255
R^2	0.441	0.024	0.065	0.020	0.346	0.388
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The third panel regression model regresses daily absolute return gaps (|ETF–NAV|, |NAV–IND| and |ETF–IND|) on ETF illiquidity and other controls. This is done for both the full sample period, as well as the COVID-19 period. There are several interesting results. Firstly, we find that the main independent variable log ETF illiquidity is highly significant and positive for the ETF versus NAV tracking errors, as well as the ETF versus index tracking errors; (ß = 0.001, t = 5.370) for |ETF–NAV| and (ß = 0.001, t = 4.838) for |ETF–IND|. The measure is however not significant for the |NAV–IND| tracking errors, (ß = 0.000, t = 0.516). These results imply, similarly to the results found in the standard deviation based tracking error regression model (Table 9), that illiquidity consistently increases dailiy tracking errors where ETF returns are involved. In more illiquid conditions, arbitrage becomes less effective, which allowes ETF prices to deviate from their NAVs. Furthermore, as illiquidity does not significantly impact NAV versus Index tracking on a daily basis, it is suggested that replication quality is relatively stable, irrespectable of the ETF's liquidity.

Another notable finding is that the NAV return volatility has a positive and somewhat significant effect on NAV versus Index tracking errors, and ETF versus Index tracking errors in the full sample ((ß = 0.464, t = 3.067) for |NAV–IND|, and (ß = 0.167, t = 2.080) for |ETF–IND|). When the underlying assets of an ETF fluctuate more, i.e. have a high NAV return volatility, it becomes more difficult to accurately replicate the benchmark index. This leads to an increase in NAV versus Index tracking errors, and thus the positive coefficients are expected. Interestingly, the coefficient for the NAV return volatility is not significant for the ETF versus NAV tracking error. This implies that ETF versus NAV return differences are less sensitive to NAV volatility, and could be a result of AP activity keeping ETF prices in line with their NAVs, even when NAVs fluctuate. This interpretation is supported by Figure 3, Panel A, which shows a strong linear relationship between ETF and NAV returns (R^2 =60.6%) and a coefficient of 0.85. The tight clustering of observations around the fitted line indicates that ETF prices tend to follow NAV returns closely, even when NAVs are volatile. This pattern is aligned with the role of arbitrage mechanisms in correcting ETF mispricings and may hence explain why the ETF versus NAV tracking errors appear relatively insensitive to NAV return volatility in the regression results.

Additionally, the Index return volatility is negative and significant for the NAV versus Index tracking errors (ß = -0.432, t = -2.830). This result is in line with the other models; when Index return volatility is low, NAVs may not fully adjust to benchmark movements, which increases tracking errors.

Another finding is that fund size (log AUM) is significant only for the NAV versus index tracking error (ß = 0.000, t = 2.587). Although the coefficient is very small, the result suggest that larger funds may experience slightly higher NAV versus Index deviations. This result is in line with the coefficient in Table 9, and also corresponds with Tsalikis and Papadopoulos (2019), who argue that fund size is one of the factors with the highest influence on tracking errors. The model also reveal that the ETF Type dummies are mostly insignificant in the full sample, indicating that daily tracking errors are not strongly driven by ETF type.

Examining the COVID-19 subsample period, we see that the log ETF illiquidity measure remains positive and strongly significant for the ETF versus NAV tracking errors, as well as the ETF versus Index tracking errors; (ß = 0.001, t = 3.692) for |ETF–NAV| and (ß = 0.001, t = 3.134) for |ETF–IND|. The measure remains insignificant for the |NAV–IND| tracking errors. This confirms that ETF mispricings are still sensitive to liquidity levels, even during crises.

It is worth noting that the coefficients in this panel regression model are considerably smaller in magnitude compared to those in the annual tracking error regressions. This difference is primarily driven by the frequency of the dependent variable. While previous models examined annual tracking errors, this specification uses daily absolute return differences. Thus, the tracking errors in this model reflect short-term tracking deviations, which are typically much smaller in scale than the annual tracking error measures, that instead capture accumulated or average deviations over time, and hence appear larger in magnitude.

Table 11

ETF Illiquidity and Daily Absolute Return Difference-based Tracking Errors

Table 11 *presents coefficient estimates for daily regressions of ETF tracking errors on ETF illiquidity. The dependent variable is the daily tracking error, which is calculated by taking the absolute value of the return differences between two return types. This approach is applied to the three return pairs: ETF returns versus NAV returns (Columns 1 and 4), NAV returns versus Index returns (Columns 2 and 5), as well as ETF returns versus Index returns (Columns 3 and 6). The estimations for the full sample can be seen in Columns 1 to 3, while Columns 4 to 6 show the estimation results for the COVID-19 period. The most important independent variable is the Log ETF Illiquidity, which is computed by taking the logaritm of the relative quoted half-spread, plus an error term of 10^{-4} to ensure that log(value) never is zero. Additional variables are defined in the Appendix B. All regressions include daily fixed effects, and test statistics based on double clustered standard errors at the fund and daily level are shown in parentheses. * denote significance at the 10% level, ** denote significance at the 5% level, and *** denote significance at the 1% level.*

	All Sample			**COVID-19**		
	|ETF–NAV|	**|NAV–IND|**	**|ETF–IND|**	**|ETF–NAV|**	**|NAV–IND|**	**|ETF–IND|**
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Log ETF Illiquidity	0.001***	0.000	0.001***	0.001***	−0.000	0.001***
	(5.370)	(0.516)	(4.838)	(3.692)	(−0.009)	(3.134)
NAV Return Volatility	0.259	0.464***	0.167**	0.010	0.171	0.087
	(1.532)	(3.067)	(2.080)	(0.300)	(1.654)	(0.974)
Index Return Volatility	−0.221	−0.432***	−0.129	0.025	−0.161	−0.049
	(−1.327)	(−2.830)	(−1.566)	(0.794)	(−1.551)	(−0.532)
Log(AUM)	0.000	0.000**	0.000	−0.000	0.000**	0.000
	(0.681)	(2.587)	(1.446)	(−0.071)	(2.337)	(1.012)
ETF Type: Bond	0.000	−0.000	0.000	0.000*	0.000	0.000
	(1.338)	(−0.668)	(0.994)	(1.661)	(0.041)	(1.507)
ETF Type: Equity	0.000	−0.000	0.000	0.000*	0.000	0.001**
	(1.103)	(−0.242)	(1.426)	(1.940)	(0.742)	(2.002)
ETF Type: Money Market	0.000	−0.000	0.000	0.000	−0.000	0.000
	(0.499)	(−0.411)	(0.252)	(0.351)	(−1.140)	(0.205)
Number of Observations	191645	191645	191645	45429	45429	45429
R^2	0.222	0.306	0.203	0.236	0.092	0.222
Daily Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

To ensure robustness of the results found in Table 11, we now perform a diagnostic regression using only log ETF illiquidity as our independent variable. The results, presented in Table 12, confirm that ETF illiquidity remains a statistically significant predictor of ETF return-based tracking errors (|ETF–NAV| and |ETF–IND|) across both the entire sample, but also across the COVID-19 subperiod. As coefficients are consistent in both magnitude and significance compared to the fully specified model, we interpret that the observed relationship between ETF illiquidity and return-difference-based tracking errors is not driven by confounding control variables. Conversely, the

|NAV–IND| tracking error is still statistically insignificant, which reinforces the conclusion that market liquidity primarily affects ETF pricing and not index replication. Furthermore, the R^2 values are lower in the simplified model. This confirms the previously stated finding that control variables improve model fit.

Table 12

ETF Illiquidity and Daily Absolute Return Difference-based Tracking Errors – Diagnostic Test

***Table 12** presents coefficient estimates for daily regressions of ETF tracking errors on ETF illiquidity. The dependent variable is the daily tracking error, which is calculated by taking the absolute value of the return differences between two return types. This approach is applied to the three return pairs: ETF returns versus NAV returns (Columns 1 and 4), NAV returns versus Index returns (Columns 2 and 5), as well as ETF returns versus Index returns (Columns 3 and 6). The estimations for the full sample can be seen in Columns 1 to 3, while Columns 4 to 6 show the estimation results for the COVID-19 period. The independent variable is the Log ETF Illiquidity, which is computed by taking the logaritm of the relative quoted half-spread, plus an error term of 10^{-4} to ensure that log(value) never is zero. All regressions include daily fixed effects, and test statistics based on double clustered standard errors at the fund and daily level are shown in parentheses. * denote significance at the 10% level, ** denote significance at the 5% level, and *** denote significance at the 1% level.*

	All Sample			**COVID-19**		
	|ETF–NAV|	**|NAV–IND|**	**|ETF–IND|**	**|ETF–NAV|**	**|NAV–IND|**	**|ETF–IND|**
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Log ETF Illiquidity	0.001***	−0.000	0.000**	0.001***	−0.000	0.001***
	(5.500)	(−0.851)	(2.334)	(4.066)	(−0.172)	(3.471)
Number of Observations	290330	290330	290330	47632	47632	47632
R^2	0.202	0.017	0.022	0.221	0.04	0.208
Daily Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The regression models collectively indicate that ETF illiquidity is a consistent determinant of ETF tracking errors involving ETF returns, especially between ETF returns and their NAVs (ETF versus NAV), at both the daily and the yearly level. Across both the standard deviation based and daily absolute return difference models, illiquidity significantly increases ETF versus NAV and ETF versus index tracking errors. However, in the regression-based tracking error model, the effect of liquidity is significant only for ETF versus NAV and not for ETF versus Index tracking errors. Since the regression-based tracking error variable aims to capture persistent tracking inefficiencies over the sample, this result implies that the co-movement between ETF and index returns is less affected by market liquidity, compared to the ETF versus NAV alignment. As a result, we note that ETF pricing efficiency is most vulnerable to liquidity issues in the short run. The diagnostic tests, which excludes control variables, confirm the robustness of the ETF versus NAV and illiquidity relationship, since coefficients remain stable and significant.

Overall, NAV versus Index tracking errors remain largely unaffected by ETF illiquidity in all model specifications, implying that replication deviations between the NAV and the underlying index most likely are affected by factors other than trading frictions. Control variables, especially NAV and Index return volatility, are more relevant in explaining NAV versus index tracking errors.

6 Supplementary Robustness Checks

We apply a random-shuffle placebo test to the annual standard deviation-based tracking errors (Table 9) and the daily absolute return difference-based tracking errors (Table 11). By randomizing each ETF's illiquidity series across time, we break any true relationship between the lagged liquidity (t-1) and the next period tracking error. In this randomization, each ETF's illiquidity values remain the same, but are randomly assigned across time. We then reintroduce the one-period time lag for illiquidity at time t-1 to predict tracking error at t.

This test is repeated 500 times, so we receive 500 independent shuffles to construct a null distribution, under which ETF illiquidity carries no predictive power for tracking errors. If the true ß falls outside the 95% null interval, below the 2.5 percentile or above the 97.5 percentile; we reject the null hypothesis and conclude that today's illiquidity has a statistically significant effect on tomorrows tracking error, and hence, the results in Table 9 and Table 11 are further solidified.

When we break the link between today's illiquidity on tomorrows tracking errors by randomly shuffling each ETF's liquidity series, the estimated ß for every tracking error measure for both Table 9 and Table 11 fall outside the null interval. This is seen in Table 13 and is visually plotted in Appendix A. The results show that the true ß from our previous regressions lie outside the 95% null interval, always above the 97.5% percentile. Hence, we can rule out that the True Illiquidity betas are driven by chance for Table 9 and Table 11. Thus, the placebo test confirms the results that have previously been shown – that there is a positive and significant effect between ETF illiquidity and daily absolute return difference-based tracking errors (Table 11), as well as for annual standard deviation-based tracking errors (Table 9).

Table 13

Random-Shuffle Placebo for Table 9 and Table 11

Table 13 *shows the results of the random-shuffle placebo test for the six tracking-error measures used in Table 9 (yearly standard deviation-based) and Table 11 (daily absolute return difference-based). The main regressors remain the same, as those used in Table 9 and Table 11, alongside the true ß coefficients. We randomly permute each ETF's illiquidity series 500 times within each ETF to break any existing relationship, and then compute a one-period lag placebo regressor and re-estimate the fixed-effects regression using the same control variables. The rows Mean Null, 2.5% Null, and 97.5% Null summarize the central 95% of the resulting placebo ß-distribution. Any results above or below these thresholds signal that today's illiquidity cannot predict tomorrows's tracking errors.*

	σ(ETF–NAV)	σ(NAV–IND)	σ(ETF–IND)	|ETF–NAV|	|NAV–IND|	|ETF–IND|
True Illiquidity ß	0.021	0.003	0.020	0.001	0.000	0.001
Mean Null	0.012	0.002	0.012	0.000	0.000	0.000
2.5% Null	0.010	0.001	0.010	0.000	0.000	0.000
97.5% Null	0.014	0.003	0.014	0.000	0.000	0.000

7 Conclusion

Mispriced and Misunderstood? This thesis aims to examine whether, and to what extent, illiquidity in Europe's fragmented ETF market is a primary driver of tracking errors in European ETFs. The findings demonstrate that ETF illiquidity significantly weakens the ETF price alignment relative to its NAVs, supporting the view that ETFs are indeed *mispriced* when arbitrage mechanisms by authorized participants fail to function efficiently. However, tracking errors between NAVs and benchmark indices seem less linked to liquidity. This indicates that certain causes of ETF performance deviations may still be *misunderstood*.

These findings shed light on a key structural challenge in the European ETF market; the fragmentation across multiple trading venues, all featuring different currencies and trading hours. This fragmentation ultimately limits liquidity and leads to sustained tracking errors. In this context, Euronext's proposed consolidation of over 3300 ETPs into one single trading venue stands out as an interesting approach to combat the problems posed by the liquidity dilution in the European market. By concentrating trading activity and improving market liquidity, such a consolidation may help reduce tracking errors in Europe. This could, as a result, enhance the alignment between ETF prices, their net asset values and their underlying indices – ultimately decreasing costs for investors.

We further suggest that future research could usefully investigate how structural changes – such as market consolidation – might improve liquidity and enhance ETF tracking accuracy in the European market. As a final remark, while *mispricing* remains a challenge, this thesis helps bring clarity into the mechanisms behind it – leaving the European ETF market perhaps a little less *misunderstood*.

References

Acharya, Viral V., and Lasse Heje Pedersen. 2005. “Asset Pricing with Liquidity Risk.” *Journal of Financial Economics* 77 (2): 375–410. https://doi.org/10.1016/j.jfineco.2004.06.007.

Ackert, Lucy F., and Yisong S. Tian. 2008. “Arbitrage, Liquidity, and the Valuation of Exchange Traded Funds.” *Financial Markets, Institutions & Instruments* 17 (5): 331–62. https://doi.org/10.1111/j.1468-0416.2008.00144.x.

Amihud, Yakov, and Haim Mendelson. 1986. “Asset Pricing and the Bid-Ask Spread.” *Journal of Financial Economics* 17 (2): 223–49. https://doi.org/10.1016/0304-405X(86)90065-6.

Bae, Kyounghun, and Daejin Kim. 2020. “Liquidity Risk and Exchange-Traded Fund Returns, Variances, and Tracking Errors.” *Journal of Financial Economics* 138 (1): 222–53. https://doi.org/10.1016/j.jfineco.2019.02.012.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi. 2018. “Do ETFs Increase Volatility?” *The Journal of Finance* 73 (6): 2471–2535. https://doi.org/10.1111/jofi.12727.

Blackrock. 2021. “Different Types of ETFs.”. https://www.blackrock.com/sg/en/ishares/education/types-of-etfs.

Blackrock. 2023. “Authorised Participants and Market Makers – Part 1.” https://www.blackrock.com/au/insights/ishares/authorised-participants-and-market-makers.

EY. 2016. “Integrated Innovation the Key to Sustainable Growth.” EY. https://eyfinancialservicesthoughtgallery.ie/wp-content/uploads/2016/10/Global-ETF-survey-2016.pdf.

Ferri, Richard A. 2024. “Index ETFs.” In *Wiley Global Finance*. https://www.fidelity.com/learning-center/investment-products/etf/types-of-etfs-index#:~:text=An%20index%2Dbased%20ETF%20seeks,to%20replicate%2C%20less%20the%20fees.&text=Most%20exchange%2Dtraded%20funds%20(ETFs,the%20performance%20of%20an%20index.

Grill, Michael, Claudia Lambert, Philipp Marquardt, Gibran Watfe, and Christian Weistroffer. 2018. “Counterparty and Liquidity Risks in Exchange-Traded Funds.” Financial Stability Review. European Central Bank. https://www.ecb.europa.eu/press/financial-stability-publications/fsr/special/html/ecb.fsrart201811_3.en.html.

Holden, Craig W., Stacey E. Jacobsen, and Avanidhar Subrahmanyam. 2014. “The Empirical Analysis of Liquidity.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.2402215.

Horton, Melissa. 2024. “Exchange-Traded Funds (ETFs) vs. Closed-End Funds: What’s the Difference?” In *Investopedia*. https://www.investopedia.com/ask/answers/052615/what-difference-between-exchange-traded-funds-etfs-and-closed-end-funds.asp#:~:text=ETFs%20are%20open%2Dended%20funds,the%20fund%20is%20%22closed.%22.

Johnson, Steve. 2025. “Euronext Plan to Consolidate ETF Trading Venues Sparks Scepticism.” *Financial Times*, 2025. https://www.ft.com/content/dd60a918-b8e9-4cf0-afc2-ab8697259d2a.

Liberto, Daniel. 2024. “Passive ETF: What It Is and How It Works.” In *Investopedia*. https://www.investopedia.com/terms/p/passive-etf.asp.

McWhinney, James. 2024. “Active vs. Passive ETF Investing: What’s the Difference?” In *Investopedia*. https://www.investopedia.com/articles/mutualfund/05/activepassive.asp.

Pástor, Ľuboš, and Robert F. Stambaugh. 2003. “Liquidity Risk and Expected Stock Returns.” *Journal of Political Economy* 111 (3): 642–85. https://doi.org/10.1086/374184.

PwC. 2022. “ETFs 2026: The next Big Leap.” https://www.pwc.com/gx/en/financial-services/publications/assets/ETF_2026_PwC.pdf.

Stoxx. 2025. “25 Years of ETFs in Europe: A Revolution in Portfolio Trading.” 2025. https://stoxx.com/25-years-of-etfs-in-europe-a-revolution-in-portfolio-trading/#:~:text=On%20April%2011%2C%202000%2C%20the,EUR%2017%20billion%20under%20management.
Tsalikis, George, and Simeon Papadopoulos. 2019. “ETFS – Performance, Tracking Errors and Their Determinants in Europe and the USA.” *Risk Governance and Control: Financial Markets and Institutions* 9 (4): 67–76. https://doi.org/10.22495/rgcv9i4p6.
U.S. Security and Exchange Commission. 2024. “Exchange-Traded Funds.” https://www.investor.gov/introduction-investing/investing-basics/investment-products/mutual-funds-and-exchange-traded-2.
Vanguard. 2024. “Capital Market Fragmentation and the Impact on Total Cost of Ownership for European ETF Investors.” https://www.ie.vanguard/content/dam/intl/europe/documents/capital-market-fragmentation-and-the-impact-on-total-cost-of-ownership-for-european-etf-investors-eu-en-pro.pdf.
Vette, Nander de, Benjamin Klaus, Simon Kördel, and Andrzej Sowiński. 2023. “Gauging the Interplay between Market Liquidity and Funding Liquidity.” European Central Bank. https://www.ecb.europa.eu/press/financial-stability-publications/fsr/special/html/ecb.fsrart202305_01~830184261b.en.html.

Appendix

Appendix A

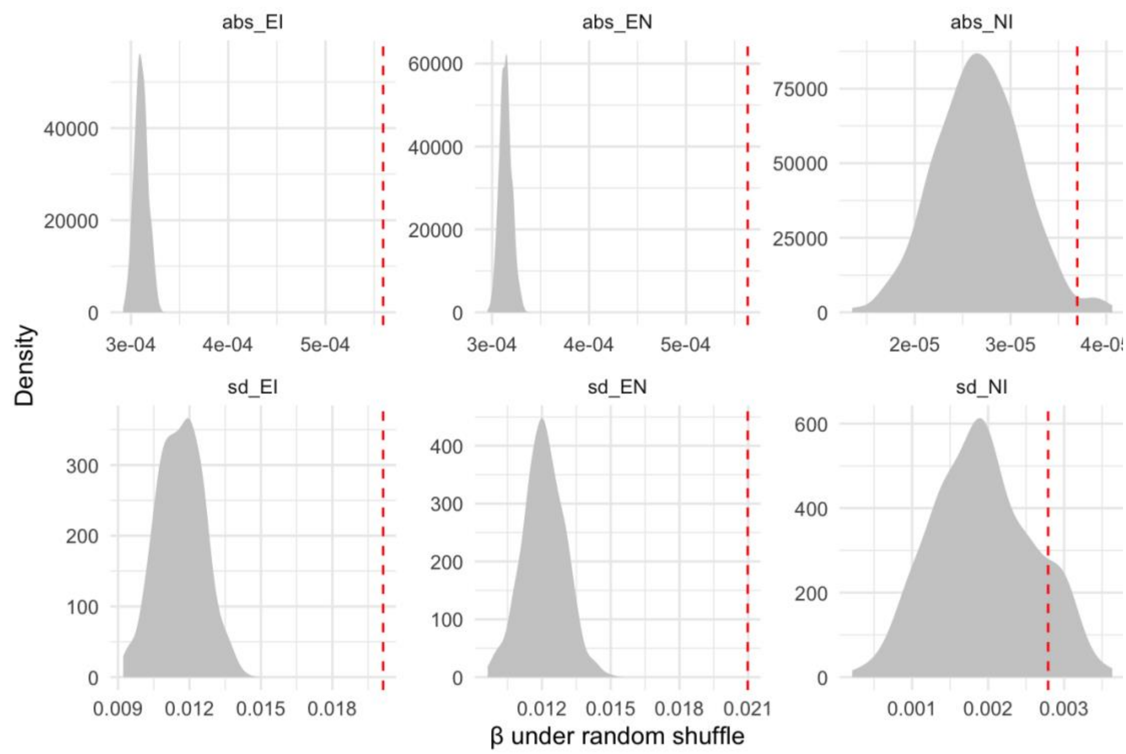
The plots in Figure 6 are merely a visualization of the results in Table 13, plotting the entire null distributions in grey for each of the six tracking errors. Although only one line is drawn for each True ß, the upper limit of the null distirubtion (97.5 percentile) is marked by the point beyond which the grey distribution thins out. The figure shows that the red dotted line for the observed relationship between today's ETF illiquidity on tomorrows tracking error is stronger than would occur by chance.

Figure 6

Null Distributions - Random-Shuffle Placebo

Figure 6 *shows our random-shuffle placebo distributions from Table 13. Each panel's grey curve is the density of 500 shuffled-illiquidity β's. The red dashed line is the true β from the original regression. The top row plots the absolute-difference errors (|ETF–NAV|, |NAV–IND|, |ETF–IND|) and the bottom row the annual σ-based errors (σ(ETF–NAV), σ(NAV–IND), σ(ETF–IND)). When the true β lies to the right (or left) of the 2.5th–97.5th percentile of the null density, it indicates that our observed illiquidity effect is highly unlikely to occur by chance.*

Random-Shuffle Placebo for Standard Deviation and Absolute Tracking Errors



Appendix B

Specification of Regression Variables

	Formula
ETF Illiquidity	$RQHS_t = \frac{P^A{}_t - P^B{}_t}{2\ x\ \hat{P}_t}$ Where: $\hat{P}_t = \frac{(P^A{}_t + P^B{}_t)}{2}$
Log ETF Illiquidity, daily	$Log(RQHS)_{i,t} = \log\ (RQHS_{i,t} + \varepsilon)$ $\varepsilon = 10^{-4}$
Log ETF Illiquidity, yearly	$Log(RQHS)_{i,t} = mean(\log\big(RQHS_{i,\tau} + \varepsilon\big))_{\tau \in t}$ $\varepsilon = 10^{-4}$
Log return ETF	$r_{i,t}^{ETF} = \log\big(CloseETF_{i,t}\big) - \log\ (CloseETF_{i,t-1})$
Log return NAV	$r_{i,t}^{NAV} = \log\big(NAV_{i,t}\big) - \log\ (NAV_{i,t-1})$
Log return index	$r_{i,t}^{IND} = \log\big(CloseIND_{i,t}\big) - \log\ (CloseIND_{i,t-1})$
NAV Return Volatility, Daily	$\lvert r_{i,t}^{NAV} \rvert$
NAV Return Volatility, Yearly	$Vol(NAV)_{i,t} = sd(r_{i,t}^{NAV})_{\tau \in t} * \sqrt{252}$
Index Return Volatility, Daily:	$\lvert r_{i,t}^{IND} \rvert$
Index Return Volatility, Yearly:	$Vol(IND)_{i,t} = sd(r_{i,t}^{IND})_{\tau \in t} * \sqrt{252}$
Log (AUM), Daily:	$Log(AUM)_{i,t} = \log\ (NAV_{i,t} * Sharesoutstanding_{i,t})$
Log (AUM), Yearly:	$Log(AUM)_{i,t} = mean\ [\log\big(AUM_{i,\tau}\big)]_{\tau \in t}$
AUM	$AUM_{i,\tau} = NAV_{i,\tau} * Sharesoutstanding_{i,\tau}$

Appendix C

AI Disclosure

AI tools were used as a supplementary aid in the coding and writing process. The LLM assisted in reviewing text for conciseness and providing grammar checks. It also supported in coding and debugging tasks in R Studio.