MASTER THESIS

2021/2022

STUDENT NAME: PARENTHOU ADRIEN

RESEARCH TOPIC

European Equity Core ETFs : Tracking Error and Performance

- TUTOR'S NAME: KAREHNKE PAUL

CONFIDENTIAL ☒ No ☐ Yes

Affidavit

ESCP Europe

I, the undersigned, do hereby state that I have not plagiarised the paper enclosed and that I am the only author of all sentences within this text. Any sentence included which was written by another author was placed within quotation marks, with explicit indication of its source. I am aware that by contravening the stated ESCP Europe rules on plagiarism, I break the recognised academic principles and I expose myself to sanctions upon which the disciplinary committee will decide.

I also confirm this work has not previously been submitted during studies prior to ESCP Europe. If this work has been written during studies conducted in parallel to my time at ESCP Europe, I must state it.

I accept full responsibility for the content of this paper.

Adrien Richard Julien PARENTHOU

June 7th, 2022

Abstract

The aim of this study is to investigate the performance of European Core ETFs and the determinants of their Tracking Error so as to help retail investors avoid sub-optimal asset allocation. We use a sample of 30 Core ETFs on European equity indexes for the period 2017-2022. Five measures of the Tracking Error are computed on the ETFs of the sample and are found significant. The performance of the ETFs in the sample is measured for both returns and tracking efficiency, and a ranking is displayed using the Information Ratio. We find the Information Ratio to be a relevant measure for retail investors. The difference between the selection of Net returns, Gross returns, and Price returns indexes is made and some advice are given to help retail investors in their asset allocation process. We find the Average Tracking Error in our sample to have grown over the past 2.5 years contrary to our initial hypothesis.

We find Size, TER, Replication Method and Volatility to have a statistically significant impact on the Tracking Error of the ETFs in our sample but find Size and TER to have surprising negative coefficients. Finally, we give some advice for the benefit of retail investors regarding simple guidance to apply when making asset allocation decisions in ETFs.

Contents

List of Tables

1 Introduction

Exchange Traded Funds are an increasingly popular investment, especially among non-institutional investors who seek to follow the market returns while diversifying at low-cost. ETF providers have been launching Core segment of ETFs to capitalize on the development of ETFs and of 'Core-satellite' investment strategies, as well as a number of specialized ETFs. As such, the amount of literature available on the subject grows as well but stays mainly aimed at U.S or Asian markets, and mainly focuses on the difference in performance between an ETF and its benchmark.

This lack of literature on European ETFs, and more precisely on Core ETFs doesn't help retail investors in their asset allocation process, that often ends up being sub-optimal with excessive investments made in dominated products as demonstrated by Ben-David & al, 2021.

As ETF are rising in popularity among investors, the study of their performance and of their tracking abilities demands attention. The goal of this paper is therefore to study the tracking efficiency and the performance of European Core ETFs. This is of interest for non-institutional investors for whom a democratized product does not necessarily mean a democratized comprehension of said product. For this purpose, we use a sample of 30 European Core Equity ETFs following broad-based indexes and study their performance, their Tracking efficiency, and the characteristics influencing the existing gap in performance between an ETF and its benchmark.

To do so, we start by putting the development of ETFs into context before looking at the existing research literature on the subject. We then define appropriate performance and tracking ability measures. The empirical research therefore focuses on the ability for Core European ETFs to track their benchmark, and on the deviations in returns experienced by the funds in the sample. To do so, the Tracking Error of those funds is computed using four different measurements, and a Student's test is performed to study the evolution of the average tracking error over time. A number of variables are then regressed on those Tracking Error measurements to test their statistical significance and add to the existing literature on the subject.

This study thus adds to the existing research by providing evidence about the existence of a Tracking Error for European Core ETFs, and by studying the explanatory variables of said tracking error, as well as in providing useful guidance intended for retail investors. The main research questions are therefore the following : Do European Core ETFs experience Tracking Error and how has it evolved over time ? What are the determinants of said Tracking Error in this context ?

The outline of this study is detailed at the end of the following Section introducing ETFs.

2 Characteristics of ETFs

The turmoil in the markets caused by Coronavirus concerns over the past 2 years has not discouraged investors from investing in Index Funds, and especially in Exchange-Traded-Funds (ETFs). With more than \$1.14tn invested in ETFs in 2021 as of the end of November, more than the record \$762.8bn gathered in 2020[1], global ETFs assets under management (AUM) reached \$10tn in December. This number is growing exponentially, as AUM in ETFs went from \$16bn in 1998[2] to \$1.41tn in 2010 to over \$10tn today. This massive amount of AUM in ETFs illustrates how important ETFs are in today's investment management business.

To properly understand why this class of financial product is so popular, one has to understand precisely what ETFs are. An exchange-traded fund is an investment vehicle first created in the 1990s for both retail and institutional investors. They "offer investors a way to pool their money in funds that makes investments in stocks, bonds, or other assets and, in return, to receive an interest in that investment pool" (SEC, 2012). Hill & Al (2016) argue that ETFs provide liquid access to practically every corner of the financial markets, enabling institutional-level portfolios to be built for a fraction of the cost of traditional mutual funds, as they are passively managed. Their main objective is therefore to track as precisely as possible the performance of the benchmark they follow, an underlying (customized) index. The difference in performance between an ETF and its benchmark is almost always negative for several reasons: ETFs providers charge management fees, transaction costs and trustee expenses can occur (see Blackrock, 2016). This difference in performance is called the Tracking Error (TE). ETFs differ from mutual funds in the sense that their shares are traded on exchanges, that they have very short

[1] Chris Flood, C., 2021. "ETF assets close to \$10tn after second year of record growth", Financial Times
[2] Investment Company Institute, A close look at ETF households, September 2018

time-to-market and that they are better suited to meet investor demand for "trendy" investing topics because of their intraday liquidity.

Novick (2017) claims that the broad and increasing adoption of ETFs is often seen as the democratization of investments. According to this view, even individual investors can use ETFs to achieve low-cost portfolio diversification along with exposure to a broad range of investment styles without resorting to expensive asset managers. However, the recent evolution of the ETF industry into two separate paths to answer investors needs puts this narrative in perspective. Indeed, while broad-based, "Core" ETFs provide investors with an opportunity for low-cost diversification, more specialized ETFs appear to grab investors' attention for trendy, yet often overvalued, investment topics, as showed by Ben-David, Franzoni, Kim & Moussawi (2021). Those more recent specialized ETFs often underperform in terms of risk-adjusted returns after fees but are still abnormally overrepresented. Brown, Cederburg & Towner (2021) show that a large portion of these dominated ETFs claim unique strategies despite their high correlations with cheaper alternatives, and that investors are allocating excessively large amounts to these funds. Such specialized ETFs cover an increasingly broader array of investments, from trade war, vegan products, bitcoin, work from home to Covid-19 vaccines, further moving away from their initial purpose to track the performance of specific equity indexes. If index-based ETFs are still predominant, newer actively managed ETFs often seek to reach a specific investment objective (SEC, 2012). Those two segments of the ETF industry are competing on different dimension: broad-based ETFs compete on price, while specialized ETFs compete on quality, understood as the product attributes other than price that attract investors. As a consequence, some investors might answer to investor sentiment (see Lee & Al, 1991), defined by Ben-David & Al (2021) as expectations of investors on the returns of an asset that are not based on fundamentals, in buying those newly ETFs tailored to appeal to investor's irrational beliefs

On the other hand, broad-based "Core" ETFs competing on prices have seen their fees dropping largely due to increased competition in the ETF space. Indeed, according to calculations by JPMorgan, asset-weighted annual fees for US-listed ETFs dropped 43% to reach 19 basis points (bps) between 2012 and 2020[3], and the same trend could be observed in Europe, Canada and Asia, with asset-weighted annual fees dropping by 51% over the same period.

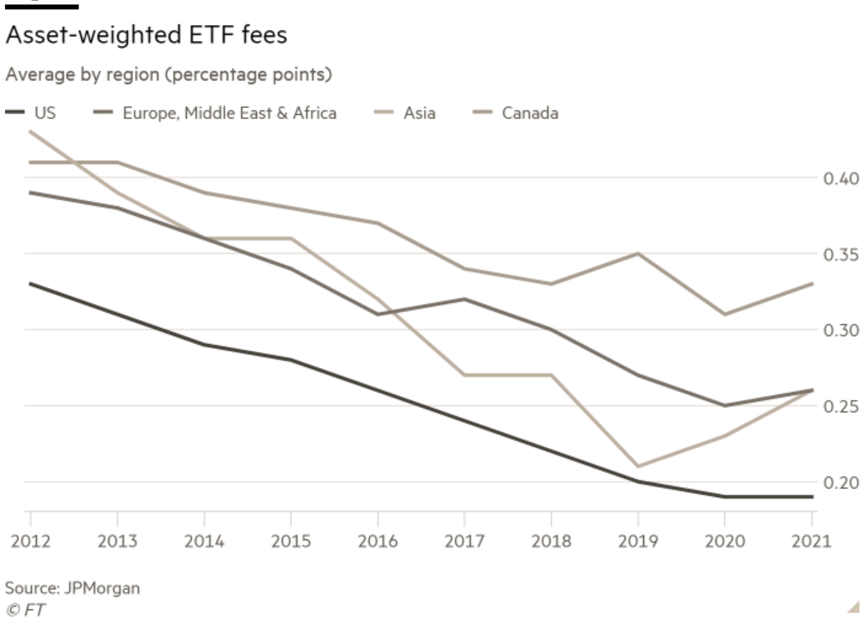


Figure 1. Asset-weighted ETF fees
Source: JPMorgan © FT

A consequence of this dualistic segmentation of the ETF market led to the application of the "Core-Satellite" approach to the sector of ETFs. The Core-Satellite approach divides a portfolio into a Core part, designed to replicate the performance of the investor's chosen benchmark; and a performance-seeking part composed of one or more 'satellites', which is allowed ampler tracking error in the search of higher performance as defined by Amenc, Goltz & Grigoriu, (2010). The authors believe that the efficient use of this strategy allows investors to create portfolios that surpass broad indexes in terms of risk-returns, as well as provides investors with an access to the outperformance of the satellite while regulating the risk of underperformance through

[3] Steve Johnson (2021), "Historic trend reverses as ETF fees head higher", *Financial Times*

the core. This view has recently been brought up to date by the biggest ETFs providers such as Vanguard and Blackrock. In its more recent update of its Guide to Core-Satellite investing, Vanguard insists on the importance of asset allocation, and on how this strategy combines the "lower cost, broader diversification, tax efficiency and lower volatility" of index funds, with the "potential outperformance" of other direct investments (Vanguard, 2021). Both Vanguard and Blackrock launched their "Core" segment to provide investors with the ability to easily take on this strategy : Vanguard Core ETFs and iShares Core ETFs. This trend was also followed by European ETF providers such as Lyxor with its Core segment, Amundi with its 'Prime' segment or BNP Paribas with its 'Easy' segment, among others.

If ETFs providers seem to emphasize the performance of those Core ETFs and their ability to track the performance of their benchmark, the availability of those products to non-institutional investors makes them worth studying, to provide investors with the key to make informed decisions regarding their choices in terms of asset allocation, especially given the overrepresentation of dominated products in the ETF sector shown by Brown & al, (2021).

A measure worth studying when trying to compare ETFs is the Information Ratio. This ratio help identifying how much a fund has exceeded its benchmark over time. This measure helps investor in their asset allocation process when focusing on overperformance rather than solely on perfect replication of the benchmark.

As defined earlier, another measure of interest when comparing the performance of an ETF to its benchmark is the Tracking Error. This measure is only partly explained by the expenses and fees mentioned earlier, and additional variables need to be considered to explain why an ETF doesn't perfectly replicate its benchmark performance. The fundamental reason is that most ETFs do not physically perfectly replicate their benchmarks. Most ETFs use a different method of replication, for a variety of reasons including transaction costs and availability of securities. Therefore, several replication strategies exist.

The first one, the full physical replication of the index, is rare and mainly focused on indexes with a low number of securities, as this sort of ETF is exposed to the securities of its benchmark by purchasing them in the same proportion as the index and is therefore exposed to transaction costs that would be too high on indexes with a lot of securities.

The second one, the physical sampling of the index, is the most common replication method. With this strategy, the ETF holds a selection of securities, a "sample" of the original index, instead of replicating it 1:1. This strategy is also called the optimized methodology, in the sense that it is an "optimized subset of index securities".

The third replication method is the synthetic one. Synthetic ETFs were popular when they were first launched in the early 2000s' but dropped a lot in popularity over the last decade, and only accounted for less than 11% of all ETFs in 2020[4]. This type of replication is different from the others as here, the ETF doesn't aim to replicate the securities in the index but rather directly the performance of the index. To do so, the synthetic ETF enters a swap contract with a financial institution that promises to deliver the performance of the index in exchange for a fee. To hedge the counterparty risk, the swap is collateralized with a basket of securities to which it has access in case of default, even though those securities usually differ from the ones of their benchmark (Nasdaq, 2021).

[4] Sin,R., Psarofagis, A., 2021. Stigma surrounding synthetic ETFs should be put to rest for good. Bloomberg Intelligence

Replication methods of ETFs in comparison

	Physical	Physical (Sampling)	Synthetical
Replication method	Full replication	Sampling	Swap based
Description	The index is replicated 1:1	The ETF holds a selection of securities	The index replicates the index by using a financial derivative (swap)
Underlyings	Equities, Bonds	Equities, Bonds	Equities, Bonds, Commodities, Money Market (EONIA etc.), Short and Leverage indices
Typical characteristics of index components	Liquid securities	Illiquid securities	Liquid and illiquid securities, investment restrictions (trade restrictions, taxation), different time zones
Typical number of securities in the index	Low	High	Low to high
Sample indices	FTSE 100, Eurostoxx 50, Dow Jones 30	MSCI World, MSCI Emerging Markets	MSCI World, MSCI Emerging Markets, Eurostoxx 50, Commodities indices, Short FTSE 100, Leveraged FTSE 100

Figure 2. Replication methods of ETFs
Source : justetf.com

Although these replications strategies are designed to minimize the Tracking Error, most studies conclude that tracking error is still statistically significant in ETF replication and almost always negative (see: Tang., Xu., 2014; Elia;, 2011; Gastineau., 2004). One could wonder why an asset not perfectly replicating its benchmark performance would attract an exponentially increasing number of investors. Part of the answer is usually marketed by ETF providers : "Easy to trade, hard to beat on fees" (Vanguard, 2022), "Exchange traded funds (ETFs) offer diversified, low-cost and tax-efficient access to the world's investment markets." (Blackrock, 2022). ETFs appeal to non-institutional investors thanks to their simplicity, passive management, low fees, broad diversification, and intraday liquidity.

Given this increasing attractiveness for Core ETFs, it seems relevant to put some attention to the way it has been developed and marketed by providers in recent years. Indeed, as mentioned earlier, giant US ETFs providers such as Vanguard and Blackrock have launched their Core segments in the last decade and have been quickly followed by other European providers. Those providers have been changing the names of those segments for marketing reasons and to differentiate themselves, but Figure 3 presents some European providers alongside the name of their Core segment, the motto used to market them and the launch date of the oldest ETF in the segment (excluding existing ETFs requalified as part of the segment).

ETF provider	Segment	Motto	Launch date
Amundi	Prime	The new generation of low-cost ETF	March 2019
Lyxor	Core	Even lower fees, not at the cost of quality	Feb. 2018
BNP Paribas	Easy	An area of expertise : portfolio Core	Sept. 2013

Figure 3: Core segments of European ETF providers
Source: Provider prospectuses

Given the importance of ETFs in the investment market, the rise of the Core-satellite strategy, and the almost always negative Tracking Error in ETFs, this thesis aims at building on the small amount of literature existing on the performance of broad-based "Core" ETFs, to add to the research and help non-institutional investors make informed decision as well as avoid dominated products while refining their investment strategy.

This research paper will therefore be organized as follows: in the third section, a literature review will be discussed in three parts : one part will review the current state of the research and the potential existing gaps in the literature, the next one will assess the different elements at the origin of the Tracking Error and the last will discuss other performance measures of interest when comparing ETFs . In the fourth section, the testable conjectures will be presented. The fifth section will describe the mathematical computation of the Tracking Error as well as the other performance measures used, before introducing the methodology of the analysis through the specification of the regressions. Section six will cover the Data selection process as well as the descriptive statistics of the selected data. In the seventh section, the Analysis of the data will be performed for each of the computation of the Tracking Error, and regressions will be performed to provide statistical significance for the identified determinants of the Tracking Error. The results will then be presented before being interpreted. Finally, the last two sections will cover the conclusions of this study and its potential limitations as well as a list of potential research opportunities, followed by a list of the references used and an Appendix.

3 Literature review

The existing research literature on ETFs and Tracking Error puts a particular focus on the relevance of the Tracking Error as a mean of evaluating the performance of passive ETFs but also mentions other performance measure that can be used to help retail investors in their asset allocation process.

Before digging further into the literature, the Tracking error needs to be differentiated from the Tracking difference. Indeed, those metrics are the two most commonly used to measure replication quality (Johnson & al., 2013). Tracking error has been defined earlier and is most often computed as the standard deviation of the difference between the returns of the ETF and its benchmark (Pope & al., 1994).The specifics of its computation are developed in Section 5.1. The Tracking Difference is defined as the simple arithmetic difference between the performance of the benchmark and the ETF's performance. The tracking error therefore measures the accuracy of the replication, whereas the tracking difference simply measures the amplitude of the difference between the returns of the ETF and of the benchmark (Johnson & al., 2013). This distinction is important as the literature doesn't always make a distinction between the two terms, which can lead to confusions. Drenovak & al. (2014) use the terminology "Tracking error" for both and argue that the tracking difference is more appropriate for active funds as it provides information on the fund over or under-performance, while the Tracking error doesn't and ranks the negative and positive active returns of the same magnitude equally. The mathematical computation of the Tracking Error will therefore be discussed in Section 5.1. to clear any potential misunderstanding regarding the methodology.

First, the state of the research and the potential existing research gaps are of interest when it comes to the study of broad-based Core ETFs, as the literature on ETF is scarce and very little study has been done on European-based ETFs, even more so on broad-based ones. Given the rise of the ETF business landscape and the development of the ‘Core-satellite’ strategy, Section 3.1 will aim at defining the current gaps in the literature that contribute to the overrepresentation of dominated products and the imperfect decision-making of investors on ETF products. It will then define how this thesis aims at providing investors with additional resources to limit this issue.

Second, determinants of the tracking error are a subject of great interest in the literature. Indeed, as factors that have a quantifiable effect on the tracking capacity of ETFs, those determinants must be considered by market participants to make informed and optimal capital allocation decisions. Managements fees, Choice of replication, Dividend payments, Volatility, Transaction costs and Size are some of the most important factors that will be reviewed in Section 3.2.

Third, several other performance measures can be used to help retail investors in their asset-allocation process. Information ratio, Sharpe ratio, and Sortino ratio can indeed help non-institutional investors differentiate ETFs based on their performance or risk-adjusted performance, to go beyond an analysis solely based on efficiency of the ETF.

3.1 Current state of research and potential research gaps

Dominated products, as defined by Brown, Cederburg and Towner (2021), are abnormally overrepresented. Ben-David & al. (2021) explain this overrepresentation, especially in the case of ETFs, by an increasing tendency from ETF providers to design products catering to investors’ irrational beliefs through “attention-grabbing themes”. To avoid straying too much from more optimal asset allocation decisions in this “clickbait” investment world, investors, and particularly non-institutional ones, need to have access to and use tools so as to make informed decisions. When it comes to ETFs, the Tracking Error as defined before is the main tool to objectively differentiate ETFs in term of divergence to their benchmark. This measure is particularly relevant for

broad-based Core ETFs that aim at providing investors with low-cost portfolio diversification by following their benchmark as closely as possible, and for whose overperformance is as undesirable as underperformance.
Core ETFs furthermore represent the most important part of the decision when following a 'Core-satellite' strategy, hence the importance for an investor seeking to follow a proper and optimal asset allocation process to rely on the existing literature studying the Tracking Error.

From the recent existing literature on the subject, most of the studies were either done on specific subsets of ETFs such as Fixed-Income ETFs or on specific geographic areas. For instance, in their 2021 '*Competition for attention in the ETF space*' paper, Ben-David & al. focus on equity US ETFs, which is also the case for Brown & al. in their 2021 '*Dominated ETFs'* paper and for Buetow & Henderson in 2012. Drenovak & al (2014) focus on European Bond ETFs, and therefore do not fill the existing gap in the literature on European Core ETFs either. Frino & Gallagher (2001) specifically focus on US equities and their paper dating from 2001 makes it less relevant given the exponential growth of the ETF market since then. Saunders & Kent (2018) study the Tracking error for International ETFs, but do not put a particular focus on Core ETFs.

This thesis therefore aims at studying broad-based Core ETFs as defined when applying a 'Core-satellite' strategy to an asset allocation decision, specifically on Equity based ETFs following a European or World Index and that are not actively managed. The specifics of the Data selection process will be developed in Section 6.1.

3.2 Determinants of the Tracking Error

As stated earlier, determinants of the tracking error have been one of the most extensively researched subjects in the literature on Tracking Error and ETFs. Indeed, as factors that have a quantifiable effect on the tracking capacity of ETFs, those determinants must be considered by market participants to make informed and optimal capital allocation decisions. Even though the interconnectivity between some of those factors makes it harder to clearly define their consequences, the following section aims at providing an overview of the main ones.

<u>Choice of replication</u>

As shown by both Johnson & al. (2013) and Charupat & Miu (2013), the replication method chosen for an ETF can have a significant impact on its performance. Indeed, Charupat & Miu (2013) state that there is usually an intrinsic trade-off between performance and Tracking Error in the choice of a replication method. A full physical replication (i.e. holding the index securities in matching weights) naturally "ensures the minimum Tracking Error" but dramatically increases the transaction costs involved, whereas holding only a representative subset of the index securities significantly lowers the transaction costs at the cost of a higher Tracking Error.

Additionally, Johnson & al. (2013) argue that the synthetic replication of an index, as defined in Section 2, is supposed to outperform the physical replication method as they don't experience a "cash drag" (i.e. a delay between the time the ETF receives a dividend and when it reinvests or distributes the proceeds) and as they are subject to less trading costs. Drenovak & al. (2014) however claim that this is only due to a difference in performance measurement methods and that physical replication is supposed to outperform synthetic replication with regards to Tracking Error as a performance measure.

Whether or not the synthetic or the physical replication method outperforms the other, the choice of replication method has therefore an impact on the performance of an ETF.

Management fees

As explained in Section 2, the Tracking Error is the result of several factors, of which management fees are among the most important ones.

First, management fees correspond to the explicit costs charged by ETF managers for managing the fund. According to Charupat & Miu, (2013) they are most commonly charged on a "daily basis by the fund issuers at annualized rates, referred to as expense ratios" based on the average daily NAVs of the ETF.

Given the definition of the Tracking Error and as shown by Osterhoff & Kaserer (2016), the higher the expense ratio of an ETF, the more likely the ETF is expected to underperform its benchmark, therefore the higher the Tracking Error.

Most of the relevant literature supports this idea (see: Elton, Gruber, Comer & Li (2002); Lin & Chou (2006); Rompotis: (2006, 2011); Agapova (2011); Elia (2012) among others), but it should be noted that Rompotis (2012) cannot verify this relationship between Tracking Error and management fees on his sample of German ETFs and that Chu (2013) finds a negative relationship between Tracking Error and management fees on his sample of Hong Kong based ETFs.

Transaction costs

Transaction costs must be differentiated from other management fees, because even though they are inherent to the replication method of an ETF, they are not a consequence of direct management by the ETF provider. Indeed, transaction costs are the costs ETFs face when they adjust the subset of securities constituting their portfolio to follow the changes in their benchmark index composition. As a consequence, they are the result of the aim of an ETF to follow its benchmark constitution. According to Chu (2013), transaction costs are "generally higher for ETFs tracking indexes that are more volatile and with underlying assets that are less liquid". Chu (2013) highlights the fact that, based on the work of Gastineau (2002), transaction costs are specific to the design and management of the index. Gastineau (2002) finds that the ETFs in his study are significantly affected by changes in the index composition, and that the changes in the index lead to a "25-30 percent annual turnover of stocks in rebalancing the portfolios, resulting in an annual transaction cost of 2-3 percent". Those transaction costs therefore undermine the performance of the ETF and amplify the Tracking Error.

Frino & Gallagher (2004) furthermore find that the Tracking Error of ETFs can be influenced by the number of share issuances/repurchases and spin-offs of the companies composing the index.

Size

Another factor theorized as being among the most important ones when studying the Tracking Error is the size of the ETF. Indeed, Chu (2013) claims that the size of an ETF, understood as its AUM, is expected to be negatively related to the Tracking Error, because of the economies of scale and the lower transaction costs that larger ETFs encounter. This belief is supported by the works of Cresson & al. (2002) that finds that "large-capitalization stock funds produce less tracking error than low-capitalization stock funds" and of Larsen & Resnick (1998) that reach the same conclusion. It should nevertheless be noted that Rowley & James (2015) find that the significance of size (proxied by AUM) is questionable as it was not significant when using quoted price returns in calculating the Tracking Error, and had a minimal effect on their model when using NAV.

Treatment and importance of dividends

Linked with the "cash drag" detailed earlier, another major factor influencing the Tracking Error is the treatment of the dividends. Indeed, for "ETFs organized as unit investment trusts" (Chu, 2013), dividends from the stocks constituting the fund are accrued in a non-interest-bearing cash account before they are distributed to the ETF holders by the providers. This is decreasing the tracking capability of the ETF as it delays the reinvestment of said dividends and therefore increases the Tracking Error. Chu (2013) says that "the higher the dividend yield, the longer the time delay, the higher the underlying index return, the more negative is the impact on the ETF returns".

This has been shown in the literature by Elton & al. (2002) as they explained that the Tracking Error of the SPDR ETFs of their sample could be explained by this accumulation of dividends in non-interest-bearing accounts, as well as by Frino & Gallagher (2001) that confirmed this effect on their sample of S&P 500 ETFs.

Volatility

Based on the effects of transaction costs amplifying the Tracking Error for volatile stocks highlighted earlier and on the works of Frino & Gallagher (2001), Qadav & Yagil (2012) and Rompotis (2011), volatility can be considered as a factor having a quantifiable effect on the Tracking Error. Indeed, Rompotis (2011) finds in his study on iShares ETFs that risk, understood as their volatility, is one of the explanatory variables that can "explain the persistence in tracking error". Qadav & Yagil (2012) find that the Tracking Error is positively correlated with the daily volatility of the ETF in their study of domestic US ETFs. Finally, Frino & Gallagher (2001) highlight that if an ETF was perfectly aligned with its index benchmark, index volatility wouldn't result in Tracking Error, ceteris paribus; but that as it is impossible for an ETF to perfectly match its benchmark, the magnitude of the Tracking Error is directly linked to the extent of volatility of the securities composing the index.

Other factors

Other factors have been studied in the relevant literature to explain the Tracking Error, such as market liquidity, portfolio Beta or sector deviation from the benchmark. This thesis will not cover them in detail as they either have a marginal effect on the Tracking Error, or the results of papers on their effects are inconclusive with some papers failing to verify what had been found in previous ones.

3.3 ETF research: Performance measures vs Efficiency measures

Measuring performance for ETFs can be confusing, especially when the term 'performance' is used without distinguishing between actual performance and efficiency. Indeed, performance can be understood as both the actual performance of the ETF compared to its benchmark i.e. the excess returns it provides its holder; as well as the efficiency with which it follows its benchmark i.e. the closeness between the returns of the ETF and those of its benchmark. Without making this distinction, it can be hard for retail investors to understand and compare the ETFs available for their asset allocation process.

As defined earlier, Tracking Error is a measure of efficiency, and an ETF performance is based on how little their returns diverge from those of its underlying index. If this measure helps retail investors in choosing a Core ETF fulfilling its role in a 'Core-Satellite' strategy, it doesn't really help investors who also want to look at the actual overperformance of the ETF and are willing to accept a higher Tracking Error in exchange for better returns.

Several measures exist to answer this need, among them three of particular interest when it comes to studying ETFs : Information Ratio, Sharpe Ratio and Sortino Ratio.

<u>Information Ratio</u>

According to Goodwin, T.H., (1998), the information ratio is a measure whose main goal is to condense into a single and easily interpretable number the mean-variance properties of an active portfolio. The main assumption for this ratio is based on the works of Markowitz, that states that the standard deviation and the mean of returns are "enough statistics to characterize an investment portfolio".

The Information Ratio can be interpreted as the average excess return per unit of volatility in excess returns. In the case of ETFs, where the fund managers have to work with a predefined Benchmark they have to follow, they can try to add value to their fund by overweighting or underweighting components of the Index while maintaining the same market risk. Simply put, the Information Ratio therefore assesses the quality of the fund manager's information (or skill) that led him to weight the components in such a way, discounted by the residual risk in the betting process.

The details of the computation of the Information Ratio are displayed in Section 5.2.

<u>Sharpe Ratio</u>

Developed by the laureate of the 1990 Nobel Prize in Economics William F. Sharpe in 1966, and from who it takes its name, the Sharpe ratio measures the profitability of a portfolio in relation to its underlying risk.

Used to compared portfolios by measuring a portfolio's returns relative to the risk-free rate and adjusting for volatility, the Sharpe ratio tells investors how much they were compensated for taking risk.

The original version of the Sharpe ratio, introduced in 1966, is based on the CAPM model and on the theory of market equilibrium. As such, and according to Goodwin, T.H., (1998), it assumed the "existence of a capital market line connecting the risk-free rate with the market portfolio". The original ratio did not distinguish between upside and downside volatility, and therefore fell short in case of non-normal distribution of returns. This limitation led to a revision of the ratio by the original author in 1994, which developed an ex-ante Sharpe ratio using benchmark returns in the numerator instead of the original Risk-free rate, making it an equivalent to the Information Ratio. The details of the computation of the original Sharpe Ratio are displayed in Section 5.2.

Sortino Ratio

Developed by Frank Sortino in the early 1980s, the Sortino ratio is a measure of risk-adjusted return of an investment that can be seen as an improvement of the Sharpe ratio.

Where the original Sharpe ratio penalized both upside and downside volatility equally, the Sortino ratio uses downside deviation rather than standard deviation (as positive upside deviation is usually welcome by investors) i.e. "only those returns falling below a user-specified target are considered risky" according to Chaudhry, A., Johnson, H. L., (2008).

The Sortino ratio is therefore used to compare portfolios and investors seek to allocate their resources in a portfolio with the highest Sortino ratio possible, as it indicates that the portfolio is operating efficiently and is not taking additional superfluous risk.

The details of the computation of the original Sortino Ratio are displayed in Section 5.2.

4 Testable conjectures

As shown in the literature review, most of the existing literature findings converge towards the existence of a significant negative tracking error in ETF replication. Given the recent development of the ETF market, the existing research on the subject starting to grow old and the development of Core segments by ETF providers, it is of interest to test whether this is also the case for broad-based European Equity ETFs.

Another interesting aspect to research is the evolution of the Tracking Error over time. Strydom., Charteris., McCullough., (2015) found the Tracking Error of their sample of ETFs tracking FTSE indexes to reduce over time. Either it be because of fund managers improving their models and reducing costs because of increased competition, or because of the increasing attractiveness of ETFs that led to an increase in their size resulting in a decrease in Tracking Error as found by Drenovak & al., (2014); studying this aspect for broad-based European ETFs can be interesting so as to advise retail investors on the attractiveness of the product.

Hence the following research question :

- **Q1: Are European "Core" ETFs still subject to tracking error in the replication of their benchmark index, and has it been reducing over the past few years as a consequence of the democratization of ETFs.**

To go further in the analysis and to help investors make informed decisions regarding their choice of a Core ETF to invest into, it can be of interest to look at the overperformance of ETFs over their benchmark, to try to understand which ETFs are overperforming compared to others. It is therefore possible to study their determinants to try to figure out a “guide to the Core ETF” that would pinpoint the most important factors to consider in making the most optimal asset allocation in a Core ETF.

Hence the following research question:

- **Q2: What influence do determinants of the Tracking Error have on Core ETFs, and what are the characteristics of the determinants of the best performing ones.**

The specifics of the methodology that will be used in answering those questions is displayed in Section 5.3.

5 Methodology

Given the definition of the Tracking error and of its main explicative variables according to the existing literature, its mathematical computation is of great interest when it comes to the analysis of its results.

The computation of other performance measures such as the Information ratio, the Sharpe ratio and the Sortino ratio will also help in ranking ETFs based on their risk-adjusted performance.

The analysis of the determinants of the Tracking Error through several regressions will finally allow for interesting interpretations and potential guidance for retail investors.

5.1 Tracking error computation

As defined previously, the Tracking Error doesn't necessarily provide information on the ETF's over or under-performance and ranks the negative and positive active returns of the same magnitude equally. It is therefore important to mention that given this definition, the over-performance of the ETF compared to its benchmark is seen as equally as undesirable as the underperformance of the same importance, which makes sense because the Tracking Error is a measure of how well the ETF follows the benchmark, not a measure of whether it performs better or not (see: Drenovak & al., 2014; Saunders & Kent, 2018). This is crucial to understand, as it might seem counterintuitive to the common non-institutional investor to choose an ETF not performing as well as another on a given period but following more precisely its benchmark than the opposite, which might be indicative of the importance that specialized "trendy" ETFs have known recently.

To calculate the Tracking Error, the existing literature either uses the change in net asset value (NAV) or the quoted price returns of the ETF. Osterhoff & Kaserer (2016) as well as Charupat & Miu (2013) state that "the deviations of the returns on

NAV of ETFs from the returns of their underlying benchmark could accumulate over time", which would significantly affect the performance of said ETFs in the long-term. On the other hand, they also assess that price deviations are usually expected to "remain within tighter bounds given the arbitrage opportunities of the daily creation/redemption process of ETFs". Buetow & Henderson (2012) further emphasize this effect in saying that as the price of ETFs is fixed on the secondary market, this price reflects both the current NAV and supply and demand for shares of said ETFs. As a consequence, they focus on variations of the market price that are driven by the market, rather than on those of the NAV that only reflects the management ability to replicate the index.

As mentioned earlier, the Tracking Error is different from the Tracking difference. Drenovak & al., (2012) refer to the Tracking Difference as the active returns of the ETF and Buetow & Henderson., (2012) compute it as the difference between the returns of the ETF and of its underlying index at the end of a given period of time (i.e. daily, weekly, monthly or annually), see Equation (1).

$$TD_{i,t} = R_{ETF_{i,t}} - R_{INDEX_{i,t}} \tag{1}$$

Where :

- $R_{ETF_{i,t}}$ is the return of the i^{th} ETF in the data sample at time t.
- $R_{INDEX_{i,t}}$ is the return of the i^{th} ETF's benchmark index at time t.

$R_{ETF_{i,t}}$ and $R_{INDEX_{i,t}}$ are computed as the ETF's or the Index's market closing prices in time t divided by its market closing price in $t-1$, minus 1. See Equation (2).

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - 1 \tag{2}$$

Where :

- $P_{i,t}$ is the price of the ETF i - or the underlying Index – at time t.
- $P_{i,t-1}$ is the price of the ETF i - or the underlying Index – at time t-1.

Frino and Gallagher (2001) as well as Charupat and Miu (2013) describe several ways to measure the Tracking Error using this formula of the Tracking Difference. Based on the work of Roll (1992), Pope & Yadav (1994) and Larsen & Resnick (1998), they identified four methods to measure the Tracking Error.

First, they compute the Tracking Error as the "monthly average absolute tracking error over n months" (TE_1) as presented in Equation (3).

$$TE_{1,ETF_i} = \frac{\sum_{t=1}^{n} \left| R_{ETF_{i,t}} - R_{INDEX_{i,t}} \right|}{n} \tag{3}$$

Where :

- n is the length of the time period considered.
- $R_{ETF_{i,t}}$ is the return of the i^{th} ETF in the data sample at time t.
- $R_{INDEX_{i,t}}$ is the return of the i^{th} ETF's benchmark index at time t.

Second, they use the root-mean-square (or standard) deviation on the ETF from that of the underlying benchmark index (TE_2), computed as :

$$TE_{2,ETF_i} = \sqrt{\frac{1}{n-1} \sum_{t=1}^{n} (R_{ETF_{i,t}} - R_{INDEX_{i,t}})^2} \tag{4}$$

Where :

- n is the length of the time period considered.
- $R_{ETF_{i,t}}$ is the return of the i^{th} ETF in the data sample at time t.
- $R_{INDEX_{i,t}}$ is the return of the i^{th} ETF's benchmark index at time t.
-

Third, they describe the standard test used in the industry for the Tracking Error which measures the standard deviation of the difference between the returns of the ETF and of its underlying benchmark Index (TE_3). It is expressed as :

$$TE_{3,ETF_i} = \sqrt{\frac{1}{n-1} \sum_{t=1}^{n} [(R_{ETF_{i,t}} - R_{INDEX_{i,t}}) - (\overline{R}_{ETF_{i,t}} - \overline{R}_{INDEX_{i,t}})]^2} \tag{5}$$

Where :

- $R_{ETF_{i,t}}$ is the return of the i^{th} ETF in the data sample at time t.
- $R_{INDEX_{i,t}}$ is the return of the i^{th} ETF's benchmark index at time t.
- $\overline{\mathrm{R}}_{\mathrm{ETF}_{i,t}}$ is the sample mean return on the ETF over the considered period.
- $\overline{\mathrm{R}}_{\mathrm{INDEX}_{i,t}}$ is the sample mean return of the underlying benchmark index over the considered time period.

Regarding this computation of TE_3, Charupat & Miu (2013) comment that, contrary to TE_1 and TE_2, TE_3 can be equal to zero, even with an ETF that is consistently overperforming or underperforming its underlying benchmark. Drenovak & al. (2012), Buetow & Henderson (2012) and Tang & Xiu (2014) use another way of computing TE_3 in their works, presented in Equation (6).

$$TE_{3,ETF_i} = \sqrt{\sigma^2_{ETF_{i,t}} + \sigma^2_{INDEX_{i,t}} - 2\sigma_{ETF_{i,t}}\sigma_{INDEX_{i,t}}\rho_{ETF_{i,t}\,,\,INDEX_{i,t}}} \tag{6}$$

Where :

- $\sigma_{\mathrm{ETF}_{i,t}}$ is the standard deviation of the returns of the i^{th} ETF in the data sample over time period t.
- $\sigma_{INDEX_{i,t}}$ is the standard deviation of the returns of the i^{th} ETF's underlying index benchmark over time period t.
- $\rho_{ETF_{i,t}\,,\,INDEX_{i,t}}$ is the correlation between the returns of the i^{th} ETF and the returns of its underlying index benchmark.

Tang & Xu (2014) develop on the pros and cons of using this standard deviation of the difference in returns between the ETF and its benchmark as a measure for the Tracking Error.

They first emphasize that as this method of measuring the Tracking Error is the most popular in the literature, it has obvious advantages. Indeed, if the ETF perfectly tracks its underlying benchmark, this measure will return zero. It also takes into account both negative and positive deviations, meaning that this method is more accurate than the mean when considering the time period of the study as well as the investment horizon of the investor.

They use the following example to illustrate it : Over a one-week time horizon, if an ETF underperforms its benchmark index by 100 bps the first day of the week but compensates that underperformance by overperforming and reversing this effect during the rest of the week, the average out - or under - performance of the ETF over the

week is zero, a number that lets the investor believe that there has been no tracking risk involved over said week with the ETF. However, if you change the time horizon of the investor to only the first day of the week, it becomes apparent that some level of Tracking risk still exists, and that this is not captured by the average.
The authors also add that this measure of the Tracking Error "penalizes large deviations from the mean", which is an advantage when studying the closeness of the performance between the ETF and its benchmark.

Tang & Xiu (2014) nonetheless highlight some of the limitations of this measure. First, they explain that this measure does not properly mirror the average deviation of the returns of an ETF from its index benchmark, as the standard deviation does not reflect the average active return of the ETF. Hence why, for instance, an ETF always underperforming its benchmark by 10 bps would have active returns of -10 bps every day, leading this method of measuring the Tracking Error to return zero, which could confuse and possibly mislead investors looking at this ETF. The authors also mention that this method of measurement requires a large number of observations to produce an accurate measure. As a way to overcome those shortcomings, they state that one has to use both this method of measuring the Tracking Error as well as the mean absolute daily relative return (see Equation (5)) to make as thorough of an analysis as possible.

Finally, Frino & Gallagher (2001) as well as Drenovak & al. (2012) and Charupat & Miu (2013) develop on the OLS approach of the Tracking Error, based on the works of Pope & Yadav (1994), Cresson & al. (2002) and Rompotis (2008).
They state that an Ordinary Least Square (OLS) regression of the returns of the ETF on the returns of the index can be used as a measure of tracking performance.

The Tracking Error can therefore be estimated through the standard deviation of the following regression :

$$R_{ETF_{i,t}} = \alpha_i + \beta_i R_{INDEX_{i,t}} + \varepsilon_{i,t} \quad (7)$$

Where :

- $R_{ETF_{i,t}}$ is the return of the i^{th} ETF in the data sample at time t, here the predicted value of the dependent variable for any given value of the independent variable $R_{INDEX_{i,t}}$.
- $R_{INDEX_{i,t}}$ is the return of the i^{th} ETF's benchmark index at time t.
- α_i is the intercept of the regression associated with the i^{th} ETF, i.e. the predicted value of $R_{ETF_{i,t}}$ when $R_{INDEX_{i,t}}$ is 0.
- β_i is the regression coefficient associated with the i^{th} ETF, i.e. how much we expect $R_{ETF_{i,t}}$ to change as $R_{INDEX_{i,t}}$ increases.
- $\varepsilon_{i,t}$ is the error term of the regression associated with the i^{th} ETF at time t, i.e. how much variation there is in the estimate of the regression coefficient.

According to Charupat & Miu (2013), the sign of the estimated value of the intercept is representative of the performance of the index benchmark.

Indeed :

- A positive estimated value of the intercept α_i will imply that the ETF outperforms its underlying index benchmark.
- A negative estimated value of the intercept α_i will imply that the ETF underperforms its underlying index benchmark.

Furthermore, they state that this method of calculating the Tracking Error is equal to the one displayed in Equation (5) if the coefficient β_i is equal to 1. If both the coefficient β_i and the intercept α_i are equal to 1 and 0 respectively, then the regression is equal to the Tracking Error displayed in Equation (4).

Drenovak & al. (2012) make some assumptions on the expected results of such a regression performed on ETFs following a passive investment strategy.

They state that in this passive investment framework:

- the intercept α is not projected to be statistically different from 0,
- the coefficient β is not projected to be statistically different from 1,
- R^2 is projected to be very high.

The authors further emphasize that those expectations are closer to the truth the more the ETF follows a full replication process, which means that the more the ETF's asset allocation deviates from the weights (i.e. the asset allocation) of the index, the higher the standard error of the OLS model and the more difference in intercept α, coefficient β, and R^2.

Drenovak & al. (2014) therefore expect ETFs straying from full replication to display:

- intercept α different from 0,
- coefficient β lower than 1,
- lower R^2.

Finally, Drenovak & al. (2014) expand on a final model to compute the Tracking Error : the cointegration approach.

Based on the works of Alexander (1999) and Alexander & Dimitriu (2004; 2005), this method is more accurate on the long-term as it considers "common long-term trends in prices" for ETFs and Indexes, whereas the correlation approach is intrinsically short-term, and therefore displays substantial instability over time. This method has been popularized amongst economists, but is still far from being an industry standard, and will therefore not be dealt with in the scope of this thesis.

5.2 Computation of other performance measures

As detailed in Section 3.3, the Information Ratio, the Sharpe Ratio and the Sortino ratio are measures that help quantify the risk-adjusted performance of ETFs. It can be interesting to describe the details of their computations as it can help casual investors understand the ratios and apply them.

Information Ratio

As defined earlier, the Information ratio is simply the ratio of the excess return of the portfolio over the standard of those excess returns.

Its computation is done in four steps :

First, the excess returns must be computed over the studied period, this is expressed as :

$$ER_t = R_{p_t} - R_{B_t} \tag{8}$$

Where :

- $ER_t$is the Excess return of the active portfolio at time t.
- R_{p_t} is the return on an active portfolio at time t.
- R_{B_t} is the return on a benchmark portfolio or index at time t.

Second, the average excess return is computed as the arithmetic average of excess returns over the studied period, express as :

$$\overline{ER} = \frac{1}{T}\sum_{t=1}^{T} ER_t \tag{9}$$

Where :

- $\overline{ER}$ is the average excess return of the active portfolio over the period ranging from t=1 to T.
- $ER_t$is the Excess return of the active portfolio at time t.

Third, the standard deviation of excess returns from the benchmark (or Tracking Error following the computation displayed in Equation (4)) is computed, expressed as:

$$\sigma_{ER} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(ER_t - \overline{ER})^2} \tag{10}$$

Where :

- $\overline{ER}$ is the average excess return of the active portfolio over the period ranging from t=1 to T.
- ER_t is the Excess return of the active portfolio at time t.

Finally, the Information ratio is computed as the ratio of the average excess return over the standard deviation of excess returns, expressed as:

$$IR = \frac{\overline{ER}}{\sigma_{ER}} \tag{11}$$

Where :

- $\overline{ER}$ is the average excess return of the active portfolio over the period ranging from t=1 to T.
- σ_{ER} is the standard deviation of excess returns over the period ranging from t=1 to T.

Sharpe Ratio

As defined earlier, the original Sharpe ratio is an ex-post performance measure based on the actual returns of the portfolio over the risk-free rate.
It is therefore computed as follows:

$$SR = \frac{R_p - R_f}{\sigma_p} \tag{12}$$

Where :

- R_p is the return of the portfolio.
- R_f is the risk-free rate of return.
- σ_p is the standard deviation of the portfolio returns.

Sortino Ratio

As mentioned before, the Sortino ratio is often seen as an improvement of the Sharpe ratio, focusing on the downside risk more than on the upside one. It is computed as follows :

$$SR = \frac{R_p - R_f}{\sigma_p} \tag{12}$$

Where :

- R_p is the return of the portfolio.
- R_f is the risk-free rate of return.
- σ_p is the standard deviation of the portfolio returns.

5.3 Specification of the regressions and tests performed

To try to provide some answers to the research questions raised in Section 4, the tests and regressions that will be performed in this thesis are presented here.

After presenting the Data selection process, the Data extraction process and the issues faced along the way, the descriptive statistics of the chosen sample will be displayed in Section 6.

Section 7 will in a first part present the results of the computations of the Tracking Difference as well as of the Tracking Error for the ETFs of the sample through the four different methods presented in Equations (3) to (6) in Section 5.1. Those Tracking Errors will be averaged, and a one sample t-test will be performed to test the significance of the following null hypothesis: the mean daily tracking error is equal to zero. This will allow us to check the significance of the Tracking Error of the ETFs in the sample, and therefore test the first part of the first research question.

Section 7 will then present the computation of the Tracking error through the root-mean-square of the linear regression method presented in Equation (7) and will use the α from the OLS regressions with the ETF returns as dependent variable and the benchmark returns as independent variable to study the performance of the ETFs of the sample.

A computation of the Information Ratio will then be made for each of the ETFs and be interpreted to try to deepen the interpretation of the performance of the ETFs, and to focus the later part of the analysis on the best performing funds in the sample.

To test the evolution of the Tracking Error over time, the sample will be divided into two periods of same duration and Tracking Error will be computed for both periods. A normality test will then be performed on the samples of Tracking Errors before running a paired two-sample t-test to check if the differences are statistically significant.

Finally, Section 7 will present the data set that was constructed to control the determinants of the Tracking Error in the sample. The results of several tests that were run to choose the type of regression to be run on the data will also be presented: a F-test, a Lagrange Multiplier test and a Hausman test.

Given the results of those tests, the following fixed effect panel data regression with Tracking Error as dependent variable will be modeled :

$$TE_{k,i} = \beta_0 + \beta_1 TER_i + \beta_2 SYNT_i + \beta_3 VOL_i + \beta_4 SIZE_i + \beta_5 DIV\_Y_i + \varepsilon_i \tag{13}$$

Where :

- TER_i is the expense ratio of the fund.
- $SYNT_i$ is a dummy variable studying the replication method of the fund; it equals 1 is the fund follows a synthetic replication, 0 if it is physically replicated.
- VOL_i is a measure of volatility of the ETF measured by the standard deviation of daily returns.
- $SIZE_i$ is the natural logarithm of the monthly AUM, whose original value is in million €.
- DIV_Y_i is the monthly dividend yield of the ETF.

Simple OLS regressions with the same dependent and independent variables will be run for four measures of the Tracking Error to control the effects of the variables detailed in the previous regression as the model detailed above is too complex to put to work, even though the steps that led to this choice in the first place will be detailed to provide potential following to this study.

6 Data

The sample of 30 Equity broad-based Core ETFs used in this study is presented in table 2. The selection process used is presented in section 6.1.

6.1 ETF selection process

This thesis aims at providing European investors with useful tools to make informed investment decisions and to limit investment into dominated products when it comes to European Core ETFs. Xetra (Deutsche Börse) and Euronext are the main European stock exchanges, listing respectively over 1800 and 1400 ETFs, and as such have been the main focus in the study.

As such, the sample of ETFs studied in this thesis must respect the following criteria to be considered European Core ETFs :

- The ETF must be traded on Euronext or on Xetra
- The ETF preferably should have been listed before January 2017 and still be active as of January 2022, but exceptions have been made for 4 ETFs, 3 launched in the first half of 2018 and one launched in the first quarter of 2019.
- The ETF must track a country or a geographical area specific equity index
- The ETF cannot follow a particular investment strategy or focus on a specific subsector or theme
- The ETF currency must be EUR (€)

Applying those filters on the Euronext ETF finder (https://live.euronext.com/products/etfs/list), the Xetra ETF finder (https://www.boerse-frankfurt.de/etfs/etfs) and on a third-party website 'justETF' (www.justETF.com/fr/find-etf.html) ETF finder leads to a sample of 58 ETFs.

Out of those 58 ETFs, 30 were filtered out by applying the following criteria :

- The ETF and its underlying benchmark daily returns must be available on Bloomberg.

This results in a selection of 30 broad-based Core European equity ETFs, displayed in Table 2.

In the first column of Table 2 is presented the name of each ETF in the sample. The sample is composed of ETFs from the following providers: Amundi, HSBC, BNP Paribas, Lyxor, Vanguard and SPDR (SSGA). All the ETFs have the mention 'UCITS' in their name, which is an acronym for "Undertaking for Collective Investments in Transferable Securities" which is the equivalent of the French OPCVM "Organisme de placement collectif en valeurs mobilières". This acronym means that those funds are compliant with the UCITS regulatory framework established in the European Union and allows them to be sold in all EU countries. UCITS rules state that the funds must be liquid and must respect strict rules regarding diversification, as well as establishing a KIID "Key Investor Information Document" describing the investment strategy, the risk profile, the historical returns, the fees and the investment horizon of the fund.

The second column of Table 2 presents the underlying index benchmark that the corresponding ETF tries to replicate. The sample presents a slight overrepresentation of ETFs following Europe-wide indexes (EURO STOXX 50, STOXX EUROPE 600, MSCI Europe) compared to country specific indexes or world indexes, and an overrepresentation of developed countries indexes compared to emerging market countries indexes. This overrepresentation can be explained both by the scope of this Thesis that focuses on European equity ETFs, and by the lower (but rising) popularity of less developed countries by providers.

The third column of Table 2 displays the TER (Total Expense Ratio) of each ETF. A shown in the literature review, this fee is considered the main determinant of Tracking Error by most of the relevant literature. Several interesting remarks can be made on the TER of the selected sample. Among the ETFs selected, all have very low expense ratio , the highest being 0.35%, which reinforce the hypothesis of a trend for ETFs to compete on prices, and therefore decrease fees, because of the increased competition. Several cases of misalignment of fees for ETFs from different providers tracking the same index can be observed.

For instance, we can see in the sample that the ETFs proposed by BNP are on average more expensive than those provided by Lyxor on the same indexes: *BNP Paribas Easy EURO STOXX 50 UCITS ETF* has a TER of 0.18% whereas *Lyxor Core EURO STOXX 50 (DR) UCITS ETF Acc* has a TER of 0.07%, even though they both track the EURO STOXX 50. This is also the case for *BNP Paribas Easy STOXX Europe 600 UCITS ETF* which has a TER of 0.20% whereas *Lyxor Core STOXX Europe 600 (DR) UCITS ETF Acc* has a TER of 0.07%.

Among the providers, Lyxor has the most variance in terms of TER between the different ETFs it proposes, ranging from 0.07% to 0.35% based on the index followed and the country/geographical area of said index. Vanguard and HSBC are the two cheapest ETF providers in the sample with fees under 0.11%. The average TERs by provider are displayed in the following table, as well as the average TER for the whole sample that stands at 0.150%.

ETF provider	Average TER
Amundi	0.167%
Lyxor	0.140%
BNP	0.193%
Vanguard	0.110%
SPDR	0.215%
HSBC	0.075%
Sample	**0.150%**

Table 1: Average TERs by ETF provider

The fourth column of Table 2 presents the replication method of the ETF. The three different methods used in the sample are Physical (Full replication), Physical (Sampling) and Synthetic (Unfunded swap). We see an overrepresentation of full replication compared to what is usually the standard in the industry, as presented in Section 2, as this is the consequence of the criteria defined earlier. As shown in Section 3.2, replication method is also one of the main determinants of the Tracking Error and is supposed to be reduced when the ETF is fully replicating the index. Given the sample of ETFs, we will be able to test the difference in performance between Physical and Synthetic ETFs.

Column five of Table 2 displays the fund listing date. As Core ETFs have been more recently brought back in the spotlight because of the democratization of ETFs and of the 'Core-Satellite' approach, a 5-year time frame seems reasonable for the analysis.

In the sixth column is displayed the fund size (in millions of euros (€)) as of January 1st, 2022. The fund size varies greatly between providers, ranging from €14m to €3592m. This fund size represents the Assets Under Management retrieved from Bloomberg using the FUND_TOTAL_ASSETS function.

As size also has an impact on Tracking Error according to the literature (see Section 3.2), we will be able to study this hypothesis on the sample.

Column seven presents the distribution policy of the fund, either accumulating or distributing. An accumulating ETF automatically reinvests all dividends or interests generated by the fund, whereas a distributing ETF pays this income to the investor. As shown in Section 3.2, the magnitude and the treatment of dividends has an impact on the Tracking Error and will have to be taken into account when performing the analysis.

Finally, the last column of Table 2 shows the fund's main exchange for informative purposes.

ETF Name	Benchmark	TER	Replic. method	Fund Listing	Fund size (€M)	Distribution policy	Exchange
Lyxor Core STOXX Europe 600 (DR) - UCITS ETF	STOXX Europe 600	0.07%	Physical(FR)	03 Apr. 2013	3494	Accumulating	Euronext
Lyxor Core MSCI World (DR) UCITS ETF - Acc	MSCI World	0.12%	Physical(S)	28 Feb. 2018	1489	Accumulating	Euronext
Lyxor Core UK Equity All Cap (DR) UCITS ETF	Morningstar UK NR	0.04%	Physical(FR)	27 Feb. 2018	333	Distributing	Xetra
Lyxor Core EURO STOXX 50 (DR) - UCITS ETF	EURO STOXX 50	0.07%	Physical(FR)	03 Apr. 2013	115	Accumulating	Euronext
Lyxor Core MSCI EMU (DR) UCITS ETF - Dist	MSCI EMU	0.12%	Physical(FR)	07 Dec. 2017	327	Distributing	Euronext
Lyxor Core US Equity (DR) UCITS ETF - Dist	Morningstar US	0.04%	Physical(S)	27 Feb. 2018	56	Distributing	Xetra
Lyxor EURO STOXX 50 (DR) UCITS ETF - Acc	EURO STOXX 50	0.20%	Physical(FR)	19 Feb. 2001	3592	Accumulating	Euronext
Lyxor FTSE MIB UCITS ETF - Dist	FTSE MIB	0.35%	Physical(FR)	03 Nov. 2003	669	Distributing	Euronext
Lyxor MSCI Europe (DR) UCITS ETF - Acc	MSCI Europe	0.25%	Physical(FR)	09 Jan. 2006	950	Accumulating	Euronext
Amundi PRIME EUROPE UCITS ETF DR - EUR	Solactive GBS Developed Markets	0.05%	Physical(FR)	13 Mar. 2019	14	Distributing	Xetra
Amundi PRIME EUROZONE UCITS ETF DR - EUR	Solactive GBS Developed Markets	0.05%	Physical(FR)	13 Mar. 2019	61	Distributing	Xetra
Amundi ETF DAX UCITS ETF DR	DAX	0.10%	Physical(FR)	23 Sep. 2008	389	Accumulating	Euronext
Amundi ETF MSCI Europe ex EMU UCITS ETF EUR	MSCI Europe ex EMU	0.30%	Synthetic	15 Jan. 2010	239	Accumulating	Euronext
Amundi ETF MSCI Europe UCITS ETF DR	MSCI Europe	0.15%	Physical(FR)	29 Jun. 2016	1298	Accumulating	Euronext

Amundi ETF MSCI France UCITS ETF	MSCI France	0.25%	Synthetic	23 Sep. 2008	104	Accumulating	Euronext
Amundi ETF MSCI Spain UCITS ETF EUR	MSCI Spain	0.25%	Synthetic	23 Sep. 2008	52	Accumulating	Euronext
Amundi ETF STOXX Europe 50 UCITS ETF EUR	STOXX Europe 50	0.15%	Synthetic	29 Sep.2009	569	Accumulating	Euronext
Amundi MSCI Emerging Markets UCITS ETF EUR	MSCI Emerging Markets	0.20%	Synthetic	30 Nov. 2010	2469	Accumulating	Euronext
BNP PARIBAS EASY STOXX EUROPE 600 UCITS	STOXX Europe 600	0.20%	Synthetic	16 Sep. 2013	573	Accumulating	Euronext
BNP PARIBAS EASY S&P 500 UCITS ETF - EUR	S&P 500	0.15%	Synthetic	16 Sep. 2019	1125	Accumulating	Euronext
BNP PARIBAS EASY EURO STOXX 50 UCITS	EURO STOXX 50	0.18%	Physical(FR)	27 Jul. 2015	202	Distributing	Euronext
BNP PARIBAS EASY CAC 40 UCITS ETF	CAC 40	0.25%	Physical(FR)	7 Mar. 2005	198	Distributing	Euronext
BNP PARIBAS EASY EURO STOXX 50 UCITS ETF	EURO STOXX 50	0.18%	Physical(FR)	26 Jul. 2015	197	Accumulating	Euronext
BNP PARIBAS EASY STOXX Europe 600 UCITS ETF	STOXX Europe 600	0.20%	Synthetic	16 Sep. 2013	85	Distributing	Euronext
HSBC EURO STOXX 50 UCITS ETF EUR	EURO STOXX 50	0.05%	Physical(FR)	5 Oct. 2009	470	Distributing	Euronext
HSBC MSCI Europe UCITS ETF EUR	MSCI Europe	0.10%	Physical(FR)	1 Jun. 2010	128	Distributing	Euronext
Vanguard FTSE Dev. Europe UCITS ETF Dist	FTSE Developed Europe	0.11%	Physical(FR)	21 May 2013	2303	Distributing	Euronext
Vanguard FTSE Dev. Europe ex UK UCITS ETF Dist	FTSE Developed Europe ex UK	0.11%	Physical(FR)	30 Sep. 2014	1670	Distributing	Euronext
SPDR MSCI Europe UCITS ETF	MSCI Europe	0.25%	Physical(FR)	5 Dec. 2014	409	Accumulating	Euronext
SPDR MSCI EMU UCITS ETF	MSCI EMU	0.18%	Physical(FR)	25 Jan. 2013	236	Accumulating	Euronext

Notes: Physical(FR) replication done through Full replication; Physical(S) replication done through Sampling ; Synthetic replication through Unfunded swap.

Table 2: European Equity broad-based Core ETFs Sample

6.2 Data availability: Net returns vs Index Prices

Once the sample was selected following the criteria displayed in the previous section, the daily ETF returns were extracted from Bloomberg using the HFA function (Historical Fund Analysis).

ETFs returns were originally supposed to be extracted using the Excel API available in Bloomberg, but due to a lot of technical errors from the API with functions (such as the 1 Day Total return function 'CURRENT_TRR_1D' or the Cumulative returns 'CUMULATIVE_TOTAL_RETURNS_GROSS_DVDS' function) not working properly and the impossibility to extract data from some of the functions available on Bloomberg (such as from the TRA function or the COMP function that allows the user to compare the historical price returns, total returns and cumulative returns between an Index and its Benchmark but does not allow the data to be extracted from the Bloomberg Terminal), the choice was made in conjunction with the Bloomberg Customer Service to use the HFA function as a proxy to extract the data needed.

Even though opinions differ in the existing literature regarding the use of NAV (Net Asset Value) or daily closing prices to compute the returns of the ETF and of its underlying benchmark, the choice was made to use the NAV as it is believed to give a more accurate representation of the value of the fund as trading dynamics can influence the market price more than the NAV.

As a consequence, the MFDF function was used on Bloomberg so as to change the settings of the reported returns to NAV instead of the default setting that does it with the daily closing prices.

Another major factor regarding the extraction of the Data to perform the Analysis is the difference between Net Return Indexes, Gross Return Indexes and Price Return Indexes.

Indeed, the prospectuses of the different ETFs in the sample all list a Net Return Index as the underlying benchmark for the ETF, and never a Gross Return or a Price index. The distinction between the three needs to be made to properly understand the difference and the potential impacts on the study.

Price Return Index

A Price Return Index only reflects price changes (capital gains or losses) of the basket of securities constituting the Index. As such, it does not consider dividends, interests, rights offering and other distributions over the holding period.

Gross Return Index (or Total Return Index)

A Gross Return Index, or Total Return Index measures the index value by tracking the prices changes of the constituting securities, but considers all before tax dividends, interests, rights offering and other distribution, and assumes that all cash contributions are reinvested. This difference in consideration of the returns makes a huge difference over time when you consider the compounding of the returns of this reinvested cash. For instance, from January 1st, 2000, to January 1st, 2022, the Price Returns of the S&P 500 were of about 223.85%, while the Total Returns were of about 386.50% (extracted using the COMP function in Bloomberg).

Net Return Index (or Total Return Index)

A Net Return Index follows the same principle as the Gross Return Index in that it considers all dividends, interests, rights offerings and other distributions and assumes reinvestment of all cash contributions, but only does so after tax, which therefore decreases the impact that cash distributions reinvestment have on the overall return of the Index, but provides a potentially more accurate return measure for retail investors.

This distinction is usually not made in the existing literature and the studies cited in Section 2 and Section 3 usually only discuss the difference between NAV and price returns but take either Price or Gross Return Indexes in computing their Tracking Errors. As taking a Net Return Index usually significantly decreases the perception of the returns of the Index compared to a Gross Return one, we must anticipate that doing so might skew the computation of the Tracking Error and significantly improve the perceived performance of the ETFs compared to their Net Return Benchmark.

To try to provide an accurate view of the Tracking Error to retail investors, and because Net Return Indexes provide a 'closer-to-reality' view on the returns effectively realized by the investor than Gross Return Indexes, it was decided for this study to extract the Benchmark returns following the underlying Indexes displayed in the prospectuses of each ETF, which are Net Return Indexes for the whole sample.

A graphic example of the difference in performance between the first ETF of the sample and both its Gross and Net return benchmarks is provided in Appendix 1.

6.3 Descriptive statistics of the ETFs and Benchmarks returns

The daily returns from the ETFs in the sample and their respective underlying Net Return Benchmarks were therefore extracted from Bloomberg using the HFA function and their descriptive statistics are displayed in Table 3 and Table 4 respectively.

The data was extracted according to the following criteria:

- The daily returns were extracted from 01/01/2017 to 01/01/2022 to have enough observations in the sample.
- The returns are calculated using the NAV.
- The currency is set in Euro (€).
- The underlying benchmark selected is the Net Return Index displayed in the prospectus of the fund.

Table 3 and Table 4 present the following descriptive statistics for each of the ETF in the sample as well as for the whole sample:

- Number of observations extracted.
- Mean of the daily returns.
- Standard deviation of the daily returns.
- Minimum daily return observed during the studied period.
- Maximum daily return observed during the studied period.

Table 3 also presents the correlation coefficient of the linear relation between the ETF returns and its benchmark returns, computed as :

$$\rho_{ETF_{i,}BENCH_i} = \frac{Cov(ETF_{i,}BENCH_i)}{\sigma_{ETF_i}\sigma_{BENCH_i}} \tag{14}$$

Where :

- $Cov(ETF_{i,}BENCH_i)$ is the covariance.
- σ_{ETF_i} is the standard deviation of the returns of the ETF.
- σ_{BENCH_i} is the standard deviation of the returns of the benchmark.

Table 5 then presents the cumulative yearly returns of each ETF and their underlying index, as well as the Tracking Difference for each of the years.

Table 6 presents the overall cumulative return for each ETF, the overall Tracking Difference over the studied period and the average daily excess return (or daily mean excess return) computed as displayed in Equation (9)

ETF Name	# of Obs.	Mean	Standard Deviation	Minimum	Maximum	Correlation
Lyxor Core STOXX Europe 600 (DR) - UCITS ETF	1305	0.042%	1.116%	-11.121%	7.310%	0.9923
Lyxor Core MSCI World (DR) UCITS ETF - Acc	934	0.071%	1.228%	-8.664%	7.180%	0.7706
Lyxor Core UK Equity All Cap (DR) UCITS ETF	987	0.027%	1.196%	-10.315%	7.332%	0.8592
Lyxor Core EURO STOXX 50 (DR) - UCITS ETF	1305	0.039%	1.264%	-11.387%	8.159%	0.9969
Lyxor Core MSCI EMU (DR) UCITS ETF - Dist	1305	0.039%	1.196%	-12.003%	7.101%	0.9956
Lyxor Core US Equity (DR) UCITS ETF - Dist	987	0.101%	1.539%	-11.061%	9.281%	0.9090
Lyxor EURO STOXX 50 (DR) UCITS ETF - Acc	1305	0.038%	1.259%	-11.535%	8.217%	0.9967
Lyxor FTSE MIB UCITS ETF - Dist	1305	0.048%	1.472%	-15.672%	7.311%	0.9940
Lyxor MSCI Europe (DR) UCITS ETF - Acc	1305	0.039%	1.117%	-11.583%	6.249%	0.9907
Amundi PRIME EUROPE UCITS ETF DR - EUR	733	0.054%	1.273%	-8.897%	4.660%	0.8520
Amundi PRIME EUROZONE UCITS ETF DR - EUR	733	0.055%	1.469%	-11.611%	6.740%	0.8735
Amundi ETF DAX UCITS ETF DR	1305	0.028%	1.225%	-10.830%	8.543%	0.9916
Amundi ETF MSCI Europe ex EMU UCITS ETF EUR	1305	0.038%	1.039%	-7.843%	7.064%	0.9703
Amundi ETF MSCI Europe UCITS ETF DR	1305	0.040%	1.114%	-11.560%	7.273%	0.9939

Amundi ETF MSCI France UCITS ETF	1305	0.050%	1.303%	-13.183%	8.801%	0.9947
Amundi ETF MSCI Spain UCITS ETF EUR	1305	0.004%	1.161%	-10.578%	7.004%	0.9928
Amundi ETF STOXX Europe 50 UCITS ETF EUR	1305	0.037%	1.103%	-11.343%	5.529%	0.9892
Amundi MSCI Emerging Markets UCITS ETF EUR	1305	0.036%	1.476%	-13.046%	6.936%	0.8465
BNP PARIBAS EASY STOXX EUROPE 600 UCITS	1305	0.043%	1.147%	-11.389%	7.608%	0.9828
BNP PARIBAS EASY S&P 500 UCITS ETF - EUR	1305	0.089%	1.427%	-9.869%	8.653%	0.6386
BNP PARIBAS EASY EURO STOXX 50 UCITS	1305	0.038%	1.261%	-11.701%	8.174%	0.9948
BNP PARIBAS EASY CAC 40 UCITS ETF	1305	0.052%	1.322%	-10.928%	8.668%	0.9725
BNP PARIBAS EASY EURO STOXX 50 UCITS ETF	1305	0.038%	1.256%	-11.719%	7.026%	0.9914
BNP PARIBAS EASY STOXX Europe 600 UCITS ETF	1305	0.043%	1.117%	-11.252%	4.617%	0.9861
HSBC EURO STOXX 50 UCITS ETF EUR	1305	0.039%	1.254%	-11.890%	6.951%	0.9964
HSBC MSCI Europe UCITS ETF EUR	1305	0.040%	1.086%	-10.528%	5.335%	0.9919
Vanguard FTSE Dev. Europe UCITS ETF Dist	1305	0.041%	1.116%	-11.492%	7.174%	0.9919
Vanguard FTSE Dev. Europe ex UK UCITS ETF Dist	1305	0.049%	1.148%	-11.943%	7.363%	0.9947
SPDR MSCI Europe UCITS ETF	1305	0.038%	1.107%	-11.751%	7.067%	0.9939
SPDR MSCI EMU UCITS ETF	1305	0.040%	1.200%	-12.448%	7.156%	0.9967

Table 3: Descriptive statistics of the ETF sample

Benchmark name	# of Observations	Mean	Standard Deviation	Minimum	Maximum
STOXX Europe 600	1305	0.041%	1.100%	-11.539%	7.140%
MSCI World	934	0.070%	1.249%	-9.788%	7.390%
Morningstar UK NR	987	0.026%	1.192%	-10.358%	7.295%
EURO STOXX 50	1305	0.036%	1.237%	-11.858%	7.566%
MSCI EMU	1305	0.038%	1.178%	-12.190%	7.085%
Morningstar US Large-Mid Cap NR	987	0.101%	1.704%	-14.509%	10.431%
FTSE MIB	1305	0.047%	1.491%	-17.153%	7.852%
MSCI Europe	1305	0.039%	1.090%	-11.531%	7.192%
Solactive GBS Developed Markets Europe	1305	0.051%	1.205%	-10.540%	6.577%
Solactive GBS Developed Markets Eurozone	733	0.051%	1.299%	-11.196%	6.556%
DAX	733	0.026%	1.212%	-10.899%	8.187%
MSCI Europe ex EMU	1305	0.040%	1.049%	-10.836%	7.307%
MSCI France	1305	0.048%	1.283%	-12.784%	8.400%
MSCI Spain	1305	0.005%	1.194%	-12.574%	7.171%

STOXX Europe 50	1305	0.036%	1.081%	-11.060%	7.041%
MSCI Emerging Markets	1305	0.037%	1.210%	-8.990%	5.280%
S&P 500	1305	0.085%	1.603%	-15.182%	11.301%
CAC 40	1305	0.064%	1.399%	-14.205%	9.023%
MSCI Europe	1305	0.039%	1.090%	-11.528%	7.189%
FTSE Developed Europe	1305	0.039%	1.092%	-11.551%	7.181%
FTSE Developed Europe ex UK	1305	0.046%	1.123%	-11.896%	6.792%
MSCI EMU	1305	0.038%	1.178%	-12.190%	7.085%

Table 4: Descriptive statistics of the underlying Indexes

	2017			2018			2019			2020			2021		
ETF	ETF	BENCH	TD	ETF	BENCH	TD	ETF	BENCH	TD	ETF	BENCH	TD	ETF	BENCH	TD
1	11.00%	10.68%	0.32%	-12.25%	-12.42%	0.18%	27.48%	26.98%	0.50%	-1.91%	-2.22%	0.31%	-1.13%	-0.66%	-0.47%
2				-9.22%	-9.12%	-0.10%	27.75%	28.10%	-0.34%	7.18%	6.92%	0.27%	-0.98%	-0.46%	-0.52%
3				-5.63%	-5.43%	-0.21%	25.53%	24.22%	1.31%	-18.82%	-18.47%	-0.35%	0.00%	-1.12%	1.12%
4	10.33%	9.79%	0.55%	-13.51%	-14.24%	0.73%	28.11%	27.67%	0.44%	-2.48%	-3.42%	0.94%	-0.87%	-0.90%	0.03%
5	13.13%	13.06%	0.07%	-15.05%	-15.41%	0.36%	26.10%	25.75%	0.35%	-0.19%	-0.81%	0.62%	-1.77%	-0.87%	-0.90%
6				-0.95%	-0.98%	0.03%	34.14%	33.97%	0.17%	12.13%	12.81%	-0.68%	0.00%	-0.28%	0.28%
7	10.16%	9.79%	0.37%	-13.79%	-14.24%	0.45%	28.08%	27.67%	0.41%	-2.80%	-3.42%	0.62%	-0.84%	-0.90%	0.05%
8	17.93%	17.75%	0.18%	-17.65%	-17.70%	0.05%	33.04%	32.45%	0.59%	-5.00%	-5.13%	0.13%	-0.91%	0.00%	-0.91%
9	10.56%	10.33%	0.23%	-11.89%	-12.15%	0.26%	26.15%	26.18%	-0.03%	-3.92%	-3.84%	-0.08%	-0.41%	-0.64%	0.23%
10							15.05%	13.75%	1.30%	-3.98%	-3.89%	-0.09%	0.00%	-0.61%	0.61%
11							16.90%	13.62%	3.28%	-4.39%	-2.08%	-2.31%	0.00%	-0.83%	0.83%
12	12.60%	12.37%	0.24%	-20.86%	-21.57%	0.71%	21.66%	22.33%	-0.67%	4.47%	3.20%	1.27%	-0.59%	0.00%	-0.59%
13	7.67%	7.68%	-0.01%	-9.19%	-8.84%	-0.35%	26.29%	26.67%	-0.39%	-7.37%	-7.11%	-0.26%	-0.66%	-0.41%	-0.25%
14	10.71%	10.33%	0.38%	-12.17%	-12.15%	-0.02%	26.64%	26.18%	0.46%	-3.58%	-3.84%	0.26%	-0.50%	-0.64%	0.14%
15	14.18%	13.63%	0.55%	-10.69%	-11.16%	0.48%	31.23%	30.36%	0.87%	-5.21%	-5.04%	-0.17%	-1.21%	-1.54%	0.33%
16	12.07%	12.18%	-0.11%	-14.64%	-14.35%	-0.29%	14.97%	15.00%	-0.03%	-14.24%	-14.02%	-0.22%	-1.21%	-0.93%	-0.28%
17	9.42%	9.02%	0.39%	-11.26%	-11.33%	0.06%	27.19%	27.00%	0.19%	-7.28%	-7.47%	0.19%	-0.84%	-0.61%	-0.23%
18	21.27%	20.62%	0.64%	-13.16%	-12.86%	-0.30%	22.57%	23.34%	-0.77%	9.41%	9.68%	-0.27%	0.90%	1.17%	-0.27%
19	10.86%	10.68%	0.18%	-12.25%	-12.42%	0.17%	27.56%	26.98%	0.59%	-1.63%	-2.22%	0.59%	-0.84%	-0.66%	-0.18%
20	7.60%	7.41%	0.19%	-1.35%	-2.19%	0.85%	36.82%	35.90%	0.91%	11.32%	9.95%	1.37%	-1.31%	-0.09%	-1.21%
21	10.09%	9.79%	0.30%	-13.82%	-14.24%	0.42%	27.84%	27.67%	0.17%	-2.83%	-3.42%	0.59%	-1.00%	-0.90%	-0.10%
22	13.04%	16.40%	-3.36%	-10.92%	-9.60%	-1.32%	32.24%	37.57%	-5.33%	-6.04%	-3.77%	-2.27%	-1.47%	-1.55%	0.08%
23	10.08%	9.79%	0.30%	-13.91%	-14.24%	0.33%	27.95%	27.67%	0.29%	-2.97%	-3.42%	0.45%	-0.80%	-0.90%	0.10%
24	10.80%	10.68%	0.12%	-12.19%	-12.42%	0.23%	27.49%	26.98%	0.51%	-1.51%	-2.22%	0.71%	-1.02%	-0.66%	-0.36%
25	10.33%	9.79%	0.55%	-13.72%	-14.24%	0.52%	28.26%	27.67%	0.59%	-2.49%	-3.42%	0.93%	-0.81%	-0.90%	0.09%
26	10.50%	10.33%	0.17%	-11.92%	-12.15%	0.23%	26.50%	26.18%	0.32%	-3.59%	-3.84%	0.25%	-0.72%	-0.64%	-0.07%
27	11.18%	10.57%	0.62%	-12.17%	-12.30%	0.14%	26.56%	26.26%	0.30%	-3.00%	-3.08%	0.08%	-0.68%	-0.63%	-0.05%
28	12.61%	12.30%	0.31%	-12.83%	-13.09%	0.25%	27.92%	27.12%	0.79%	3.63%	3.04%	0.60%	-0.69%	-0.53%	-0.15%
29	9.97%	10.33%	-0.36%	-11.78%	-12.15%	0.37%	26.05%	26.18%	-0.13%	-3.70%	-3.84%	0.14%	-1.35%	-0.64%	-0.71%
30	13.55%	13.06%	0.49%	-14.80%	-15.41%	0.61%	25.87%	25.75%	0.12%	-0.16%	-0.81%	0.65%	-0.69%	-0.87%	0.17%

Table 5: Historical yearly returns of the ETFs in the sample

ETF	2017-2022 ETF	BENCH	TD	Average Daily Excess return
1	55.33%	53.19%	2.14%	0.00164%
2	66.47%	65.78%	0.69%	0.00074%
3	26.65%	25.82%	0.83%	0.00084%
4	51.41%	46.98%	4.43%	0.00339%
5	51.04%	48.98%	2.06%	0.00158%
6	100.12%	99.26%	0.87%	0.00088%
7	50.11%	46.98%	3.13%	0.00240%
8	62.20%	60.75%	1.45%	0.00111%
9	51.44%	50.34%	1.11%	0.00085%
10	39.30%	37.41%	1.89%	0.00258%
11	40.65%	37.26%	3.39%	0.00463%
12	36.18%	33.98%	2.20%	0.00169%
13	50.00%	51.83%	-1.83%	-0.00140%
14	52.69%	50.34%	2.35%	0.00180%
15	65.53%	62.90%	2.63%	0.00202%
16	5.17%	6.78%	-1.61%	-0.00124%
17	48.59%	47.30%	1.30%	0.00100%
18	47.03%	48.34%	-1.31%	-0.00100%
19	55.50%	53.19%	2.31%	0.00177%
20	116.60%	110.32%	6.27%	0.00481%
21	49.73%	46.98%	2.75%	0.00211%
22	67.60%	84.04%	-16.44%	-0.01260%
23	49.80%	46.98%	2.82%	0.00216%
24	55.53%	53.19%	2.34%	0.00179%
25	51.22%	46.98%	4.24%	0.00325%
26	52.19%	50.33%	1.85%	0.00142%
27	53.37%	51.19%	2.18%	0.00167%
28	63.45%	60.67%	2.78%	0.00213%
29	50.23%	50.34%	-0.11%	-0.00008%
30	51.92%	48.97%	2.95%	0.00226%

Table 6: Overall returns and Tracking Difference

The descriptive statistics displayed in Table 3 and 4 already give some food for thoughts regarding the performance of ETFs compared to their underlying benchmark. Indeed, we can see that for all the ETFs, the mean of their returns is apparently fairly close to the mean of the returns of their Index, as it is always less than 0.01% away. A similar observation can be made for the Standard Deviation, Minimum and Maximum of most of the ETFs on the sample that are relatively close to the ones of their benchmark.

Several outliers can nevertheless be seen, and the divergence between their descriptive statistics to those of their benchmark seem to be linked to their correlation, as the ETFs with the biggest deviation in descriptive statistics to their Index are the ones with the lowest correlation, namely:

- ETF 11: Amundi PRIME EUROZONE UCITS ETF DR - EUR with a correlation to its benchmark of 0.8735
- ETF 10: Amundi PRIME EUROPE UCITS ETF DR – EUR with a correlation to its benchmark of 0.8520
- ETF 18: Amundi MSCI Emerging Markets UCITS ETF EUR with a correlation to its benchmark of 0.8465
- ETF 3: Lyxor Core UK Equity All Cap (DR) UCITS ETF with a correlation to its benchmark of 0.8592
- ETF 2: Lyxor Core MSCI World (DR) UCITS ETF – Acc with a correlation to its benchmark of 0.7706
- ETF 20: BNP PARIBAS EASY S&P 500 UCITS ETF – EUR with a correlation of 0.6386, by far the lowest in the sample.

No clear observation can be made on the link between this lower correlation and specifics indexes, as all six of those lower correlated ETF follow different benchmarks.

Table 5 allows us to have a broader view of the yearly cumulative returns of the ETFs over the 5-year study period. Several observations can be made when looking at this table. First, we can see that the Tracking Difference is economically significant, as all ETFs show meaningful deviation in returns from those of their Benchmark. No clear interpretation can be made on the sign of this Tracking Difference, as both positive and negative Tracking Difference can be observed. This seemingly omnipresence of a Tracking Difference nonetheless supports the idea that this research and its scope are meaningful. Another observation that can be made thanks to this table is that the under or over performance of an ETF to its benchmark does not seem definite, as some of the ETFs in the sample experienced both positive and negative yearly Tracking Difference (e.g. ETF 29 that shows the following TDs for the period ranging from 2017 to 2021 : -0.36%; +0.37%; -0.13%; +0.14%; -0.71%).

Some striking under or overperformers can be observed in Table 5, and Table 6 helps summarizing this. Indeed, a big variance can be observed in the overall Tracking Difference of ETFs over the period, with ETFs 4, 11, 20 or 25 greatly overperforming their benchmarks (with average Daily excess returns of 0.00339%; 0.00463%; 0.00481% and 0.00325% respectively) while ETFs 13, 16 and 18 greatly underperformed (with average Daily excess returns of -0.00140%; -0.00124% and -0.00100% respectively). The most notable return however is the huge negative Tracking Difference of ETF 22, that performed more than 16% less than its benchmark over 5 years, for a daily average excess return of -0.01260%.

Those over or underperformances can also be observed in Figure 4, presenting the Cumulative Returns of each ETF in the sample along the Cumulative Returns of their respective benchmark. The difference in returns between ETF22 and its benchmark is also the most visually significant in this line graph.

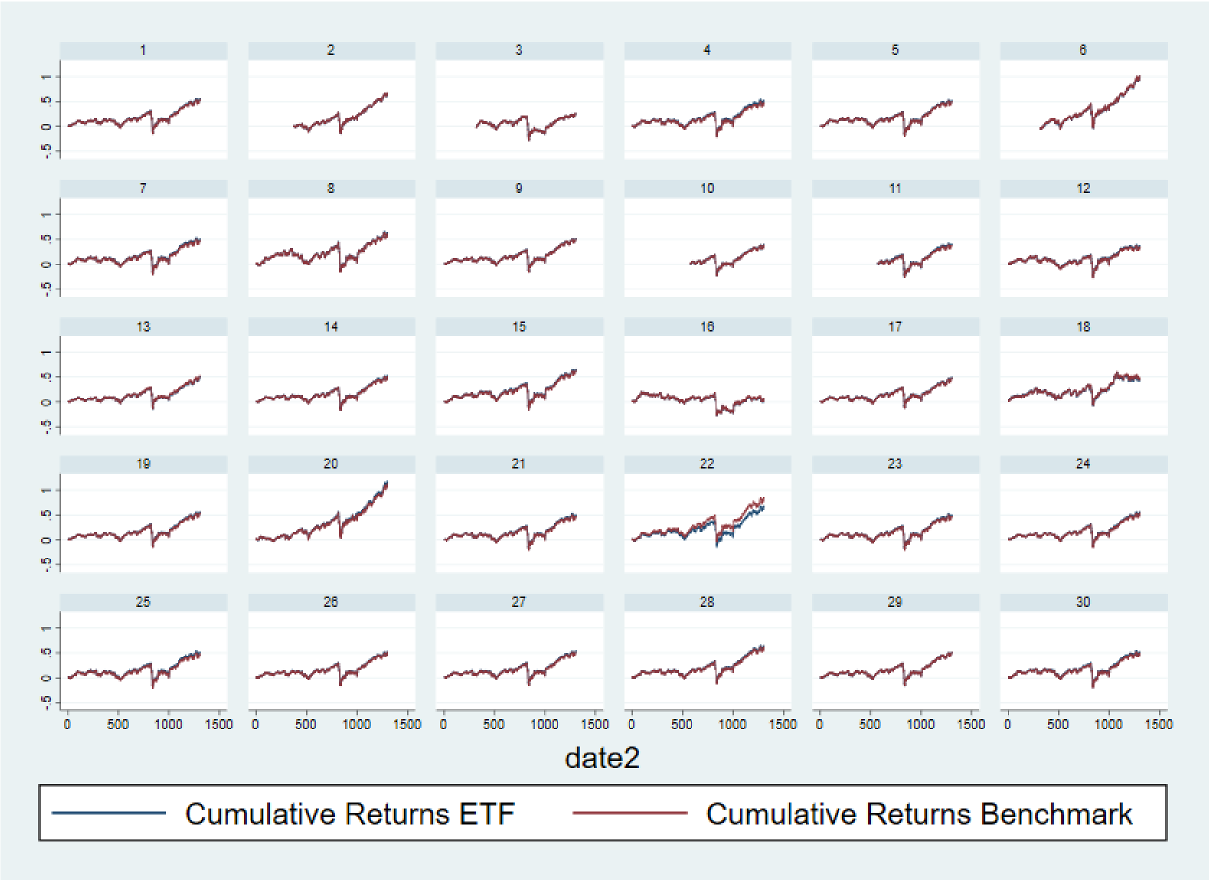


Figure 4: Cumulative Returns of the sample ETFs and underlying benchmarks

We can link some of those observations with the ones made earlier on the correlation, as the ETF with the worst correlation in the sample is interestingly the one with the highest overperformance. Out of the seven ETFs with larger Tracking Difference just listed, four are among the ones with the lowest correlations listed before. As both upside and downside variations are observed, the only thing that can be hypothesized is that lower correlation might be related to higher Tracking Difference and/or Tracking Error.

7 Tracking Error and Regression results

Given the sample of daily returns for both the ETFs and their relative benchmark, computation of the daily and monthly Tracking Difference was made, alongside computations of the Tracking Error through the measures presented in Section 5.1.

The results of the computation of the daily Tracking Difference for the whole sample are presented as a scatterplot in Figure 5.

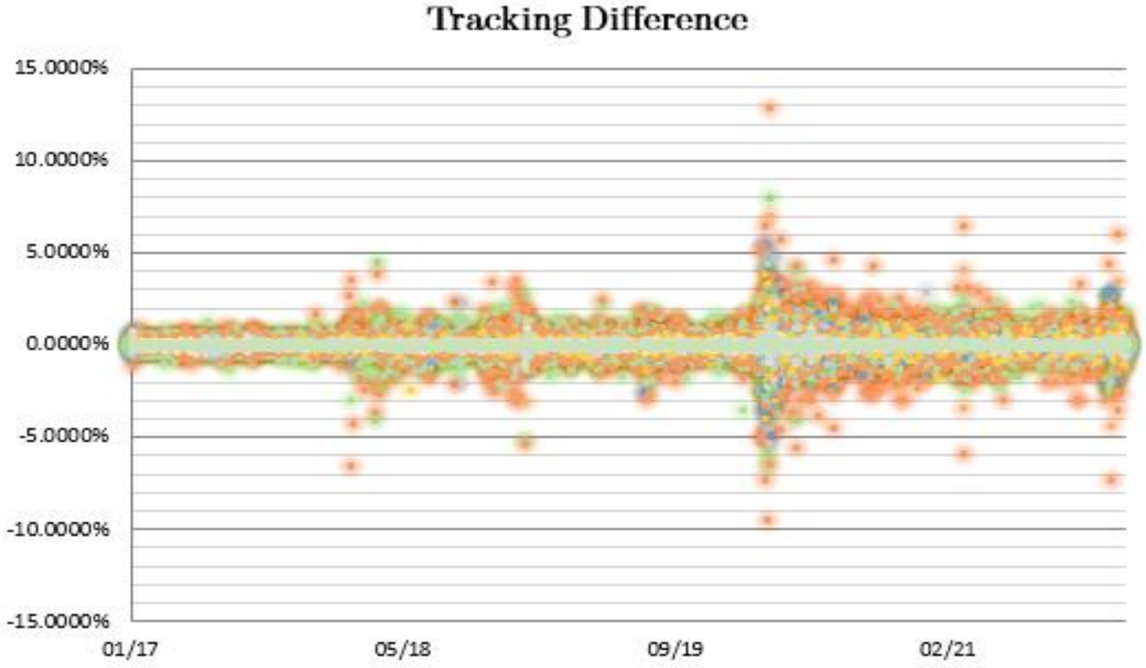


Figure 5: Full sample Daily Tracking Difference

This graphic representation of the evolution of the Tracking Difference allows for some new observations. Indeed, it appears that contrary to the hypothesis made in Section 4 that Tracking Error or Tracking Difference was reducing over time as a consequence of the increase in Size (AUM) of ETFs due to increasing asset allocation in this type of product, of the democratization of the 'Core-Satellite' strategy, and of fine-tuning of their models by ETF providers because of increased competition, the Tracking Difference seems to be increasing over time.

More than just increasing, we can see clusters of deviations from the mean in this scatterplot graph around certain periods, namely around February-March 2018, November-December 2018, and most importantly March 2020.
Those three periods can be linked to major events that happened in the stock market at those times :

- February 2018: Most major stock market indexes fell in early 2018. For instance, the Dow Jones plummeted by more than 3200 points (around 12%) in less than two weeks in early February 2018 after what is believed to have been a correction after a 20% rise in 2017 and an overall 40% rise since 2016. Fear of interest rates hikes following a Central Bank statement that noted that inflation looked to be on the rise is believed to have led to this correction.
- December 2018: As in February 2018, fear of central banks tightening monetary policies, a slowing economy and an increasing trade war between the U.S and China are believed to have led to a drop in all stock market indexes in December 2018, with the S&P 500 plunging more than 9% over the month for instance.
- March 2020: in early March 2020, most stock markets across the world crashed after increasing wariness about the spread of Covid-19. It followed reports in 2019 that the economy was slowing, decrease in manufacturing activity and geopolitical tension. The conjunction of those factors resulted in fear that the world would enter in a recession in 2020 and led to this crash.

Those three crashes having an apparently visible impact on the Tracking ability of ETFs seems to reinforce the idea that market volatility is a determinant of the Tracking Error, and furthermore that in period of high instability in the markets, ETFs struggle to properly follow the returns of their underlying index.

The full sample average daily Tracking Difference is 0.00114%, the minimum TD is -9.5351% and the maximum TD is 12.8778%, both the minimum and the maximum TD have been experienced by ETF 20: BNP PARIBAS EASY S&P 500 UCITS ETF –

EUR, which is also the ETF with the highest daily mean excess return and the lowest correlation to its benchmark. This overperformance in terms of returns but apparent under-efficiency in terms of tracking ability is a great example of the situations in which retail investors can find themselves when looking to allocate their resources. Indeed, an investor only provided with the performance of the fund would believe it is a great ETF to invest in, while another investor only provided with a measure of Tracking ability would deem it inefficient.
The Tracking Difference is furthermore a tricky measure to interpret, as averages will cancel the deviations from zero.

Further tracking measures are therefore needed and will be computed in the following parts.

7.1 TE1 - Average absolute tracking error

As defined in Section 5.1, the first computation method to calculate the Tracking Error is the average absolute TE over n periods and will be referred to as TE1.

This measure treats both negative and positive deviations equally and fits the definition of a tracking measure more precisely, as it allows the analysis of the accuracy of the Tracking rather than simple over or underperformance.

TE1 was computed using daily returns for each ETF in the sample, and Figure 6 presents a graphic representation of the evolution of TE1 over the studied period. This graphic representation helps a lot when it comes to quickly evaluate the Tracking Error of a selection of ETFs. Indeed, Figure 6 shows that ETFs 2,3,6,10,11,18 and 20 present a much larger Tracking Error than the other ETFs in the sample. ETF 20 is especially striking visually when compared to the rest of the sample.

This table also reinforces the assumptions made in the previous section, as we can see that all ETFs have seen an increase in TE1 when the March 2020 Covid crash occurred, but also puts it into perspective, as we can see that most of the ETFs experiencing a very small Tracking Error initially have faced very little increase in TE1 during this crash, whereas ETFs whose TE1 was originally higher have seen it exponentially amplified by the crash.

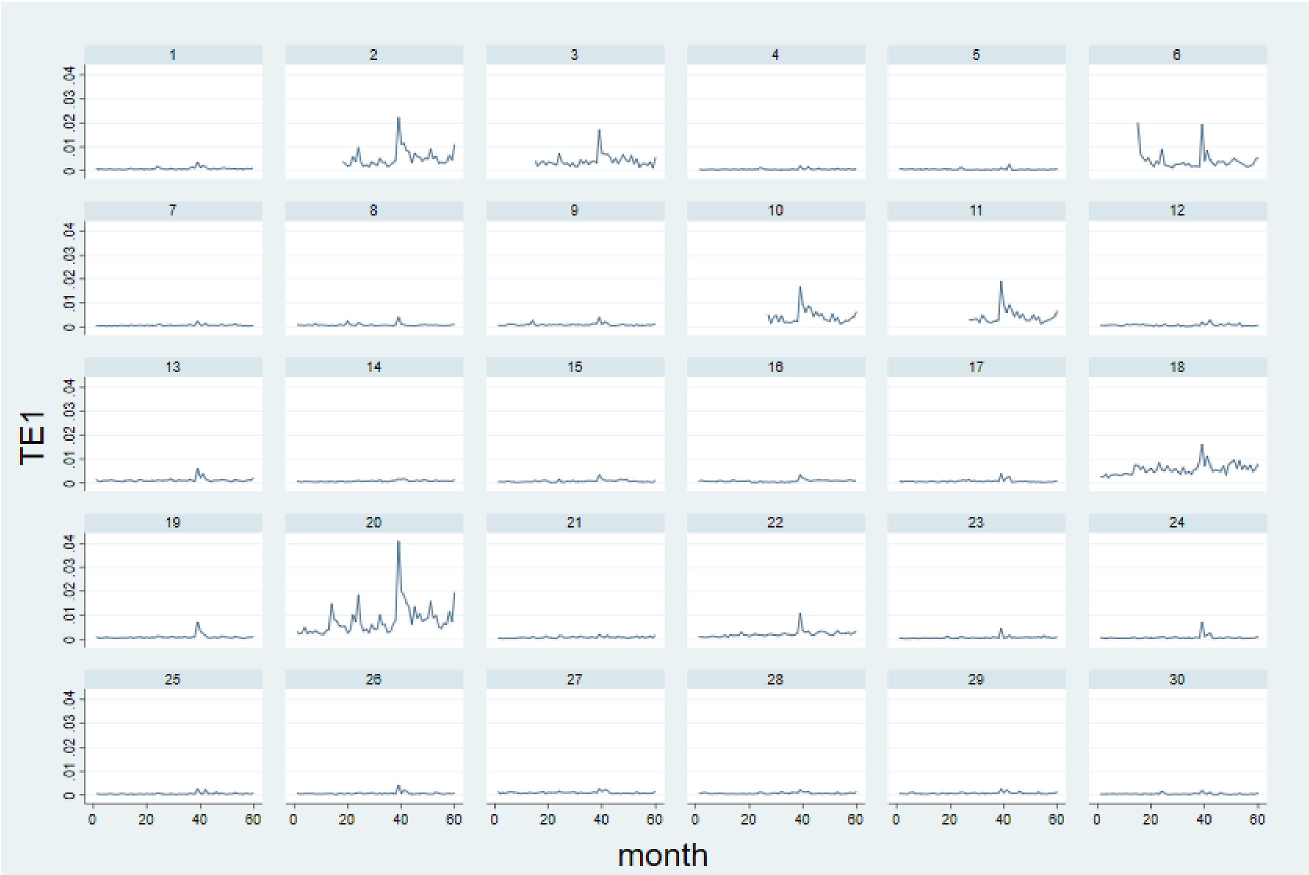


Figure 6: Tracking Error (TE1) evolution over time of all ETFs in the sample

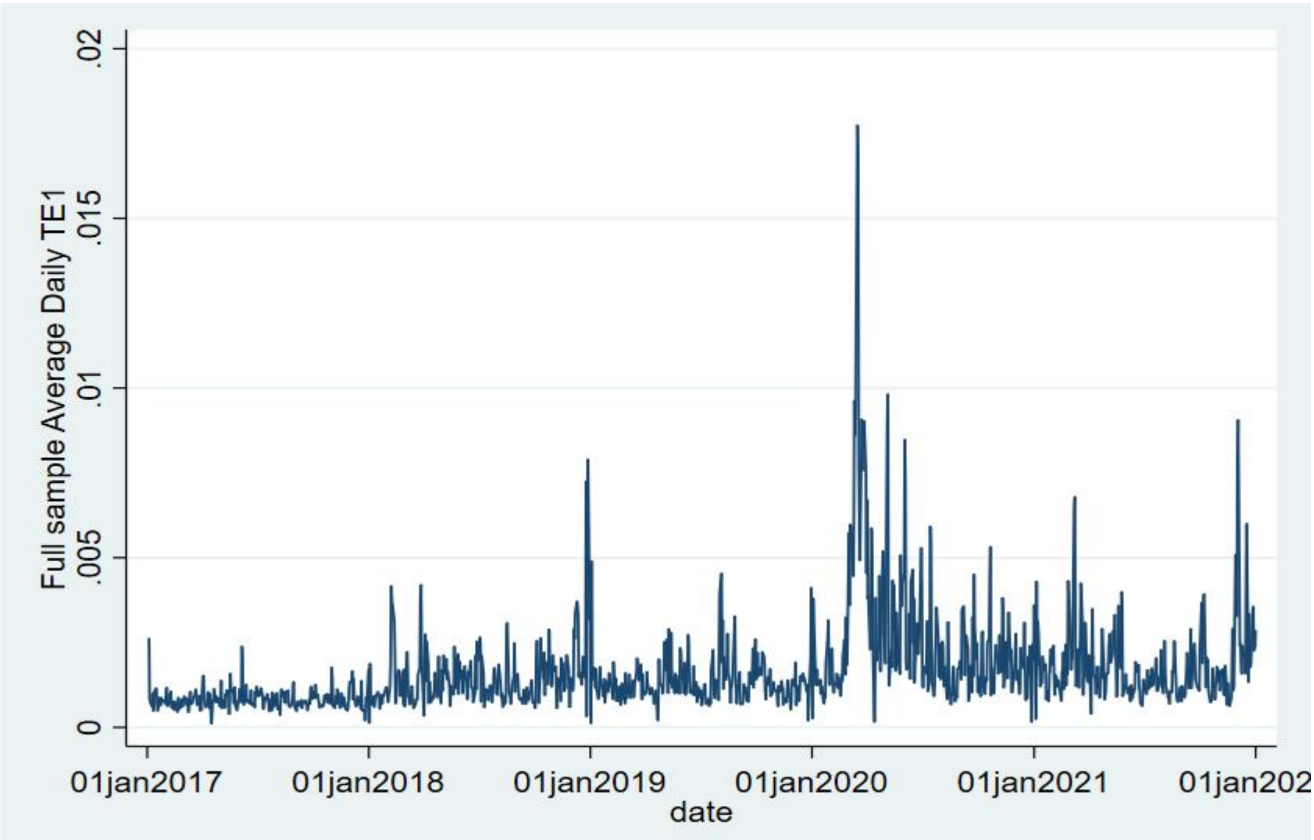


Figure 7: Sample Average Daily Tracking Error 1

Figure 7 displays the Average TE1 for the sample over the studied period, which we can link to the scatterplot of TD studied earlier. Indeed, we see spikes in Tracking Error around the three periods mentioned, and once again the Tracking Error visually seems to increase over time rather than decrease.

The average daily TE for the whole sample is 0.1646%, significantly higher than the 0.00114% of the average daily TD, which makes sense given the definition made of TE1 taking into account both upside and downside deviations. This value nevertheless stays very low, as expected when studying broad-based passive ETFs, which is a good sign for retail investors.

7.2 TE2, TE3, TE4 – Similar measures

The Tracking Error for the sample of ETFs was also computed using the computation methods displayed in Equations (4), (5) and (6) of Section 5.1.
As the values obtained using those three methods were similar at 10E-4 decimal, only the results for TE2 will be presented in this Section. All the tests or graphs performed on TE2 have nonetheless been performed for TE3 and TE4 as well and will be displayed in Appendices 2 to 5.

As for TE1, Figure 8 presents the evolution of the daily TE2 over the studied period.

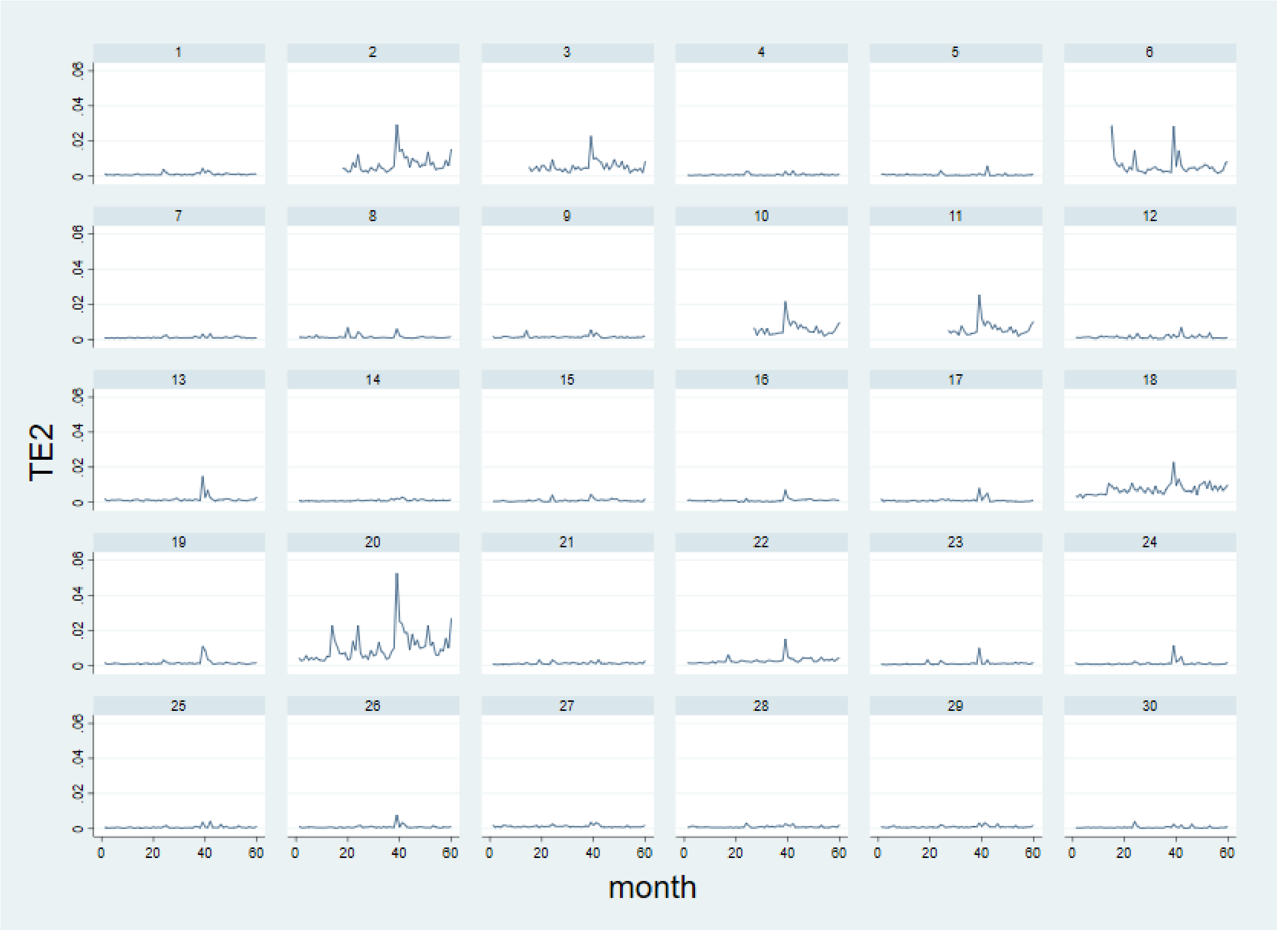


Figure 8: Tracking Error (TE2) evolution over time of all ETFs in the sample

The graphic interpretation of Figure 8 is the same as for Figure 8, the graph being very similar to the one in Figure 6. The only observation that can be made is that the ETFs with the lowest TE1 in Figure 6 see their maximum deviations slightly amplified graphically. This is due to the computation method of the Tracking Error for TE2, TE3 and TE4, that lead to higher values than of TE1 and TD.

Indeed, displayed in Table 7 are the overall (5-year) Tracking Errors, Annualized Average Tracking Difference, Daily Average Tracking difference, and overall Tracking Difference of all the ETFs in the sample, to provide an overview of the differences in scale of the values. All those values have been computed using daily data.

ETF	Avg daily TD	Annualized Avg TD	TE1	TE2	TE3	TE4
1	0.0016%	0.4275%	0.0855%	0.1380%	0.1378%	0.1380%
2	0.0007%	0.1923%	0.5311%	0.8392%	0.8390%	0.8392%
3	0.0008%	0.2198%	0.4127%	0.6338%	0.6335%	0.6338%
4	0.0034%	0.8862%	0.0580%	0.1020%	0.1020%	0.1020%
5	0.0016%	0.4112%	0.0551%	0.1128%	0.1124%	0.1128%
6	0.0009%	0.2282%	0.3907%	0.7104%	0.7080%	0.7104%
7	0.0024%	0.6251%	0.0660%	0.1031%	0.1031%	0.1031%
8	0.0011%	0.2897%	0.0825%	0.1630%	0.1623%	0.1630%
9	0.0008%	0.2209%	0.0950%	0.1527%	0.1525%	0.1526%
10	0.0026%	0.6736%	0.4338%	0.6774%	0.6773%	0.6774%
11	0.0046%	1.2112%	0.4403%	0.7154%	0.7154%	0.7154%
12	0.0017%	0.4396%	0.0756%	0.1587%	0.1587%	0.1586%
13	-0.0014%	-0.3635%	0.1222%	0.2545%	0.2546%	0.2545%
14	0.0018%	0.4699%	0.0849%	0.1240%	0.1234%	0.1240%
15	0.0020%	0.5256%	0.0809%	0.1347%	0.1347%	0.1347%
16	-0.0012%	-0.3209%	0.0863%	0.1450%	0.1444%	0.1450%
17	0.0010%	0.2591%	0.0809%	0.1620%	0.1621%	0.1620%
18	-0.0010%	-0.2597%	0.5725%	0.7865%	0.7868%	0.7865%
19	0.0018%	0.4617%	0.0969%	0.2138%	0.2138%	0.2138%
20	0.0048%	1.2579%	0.7769%	1.2977%	1.2981%	1.2977%
21	0.0021%	0.5491%	0.0811%	0.1295%	0.1294%	0.1294%
22	-0.0126%	-3.2227%	0.2075%	0.3281%	0.3279%	0.3278%
23	0.0022%	0.5642%	0.0710%	0.1650%	0.1650%	0.1650%
24	0.0018%	0.4674%	0.0767%	0.1854%	0.1855%	0.1854%
25	0.0032%	0.8480%	0.0554%	0.1065%	0.1065%	0.1064%
26	0.0014%	0.3702%	0.0708%	0.1381%	0.1381%	0.1381%
27	0.0017%	0.4354%	0.0998%	0.1428%	0.1429%	0.1428%
28	0.0021%	0.5564%	0.0759%	0.1192%	0.1192%	0.1191%
29	-0.0001%	-0.0215%	0.0779%	0.1227%	0.1217%	0.1227%
30	0.0023%	0.5889%	0.0502%	0.0985%	0.0984%	0.0984%

Table 7: Overall Tracking Errors (TE1, TE2, TE3, TE4) and Tracking Differences

As TE2, TE3 and TE4 can be qualified as standard deviation measures of returns differences, and as the Tracking Error aims at providing a measure of Tracking Efficiency, an ETF tracking its underlying benchmark perfectly would present a standard deviation of 0. As a consequence, the lower the Tracking Error measured, the better the Tracking accuracy. Figure 9 presents the Average TE2 for the sample over the studied period, which once again shows that the Tracking Efficiency of the ETFs in the sample decreases in times of financial crisis and that the Tracking Error seems to have increased over time instead of decreased.

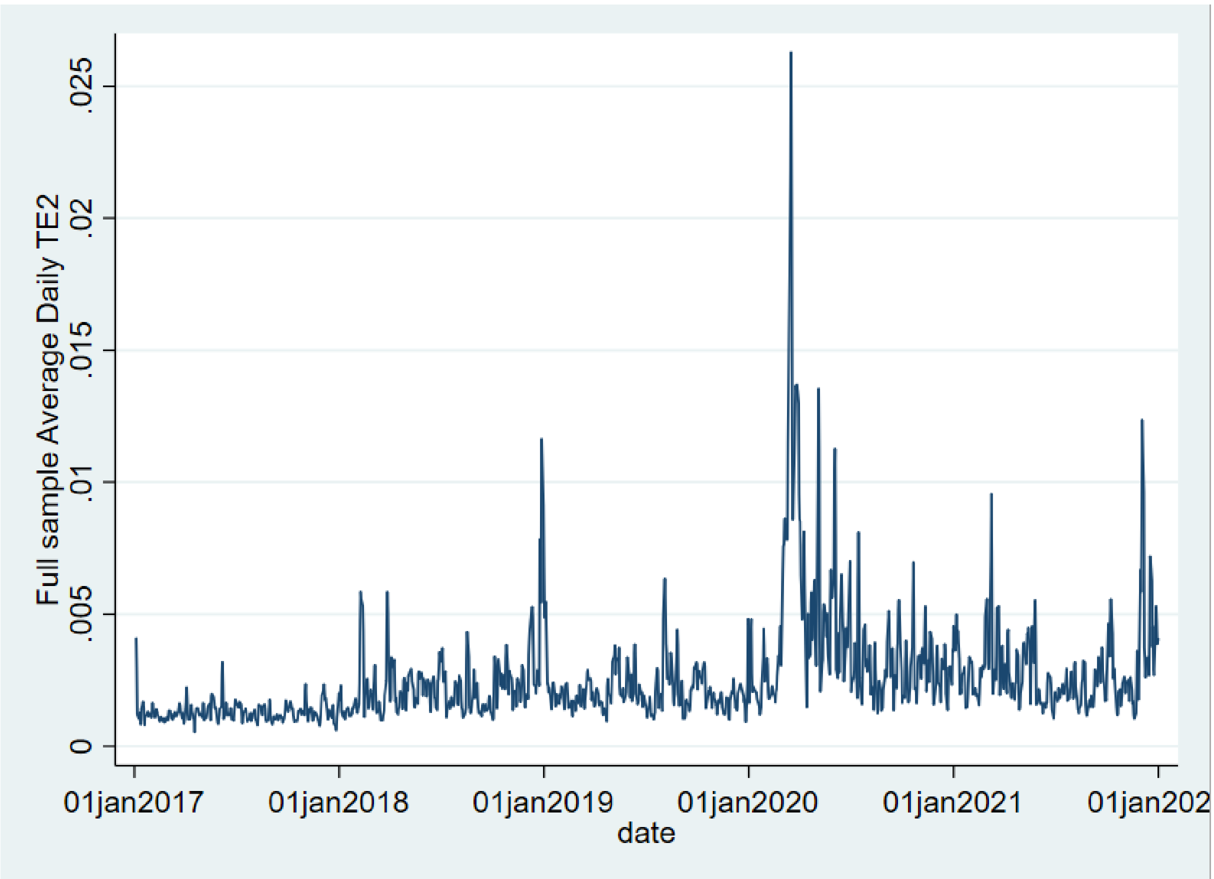


Figure 7: Sample Average Daily Tracking Error 2

7.3 TE5 and ETF performance

TE5 is computed as the standard error of the regression with the ETF returns as dependent variable and the underlying benchmark returns as independent variable, as detailed in Equation (7).

The $\propto$, β, R^2, Standard Error, t-value and statistical significance of the regressions of the daily ETF returns over the daily benchmark returns on a 5-year period are displayed in Table 8. All the regressions were done using Stata and inputting the robust command to use heteroskedasticity-robust standard errors and solve heteroskedasticity issues.

From this table, we see that the Standard Deviation of the regressions run using OLS are very close to the measures of the Tracking Error computed using TE2, TE3, and TE4 displayed in Table 7. We also observe a very high significance of all the Beta coefficients computed, with t-values ranging from 12.14 to 375.53.

On the other hand, the R-squared of all the regressions are presented in the fifth column of the table. The R-squared, or coefficient of determination, can be interpreted as the proportion of the variation in the dependent that is explained by the independent variable(s). A dependent variable perfectly explained by the independent variables introduced in the regression would therefore have a R-square equal to one, meaning that 100% of the variation in the dependent variable is explained by the independent variable(s). As such, an ETF perfectly tracking the returns of its underlying benchmark would have an R-squared equal to one. The further from one the R-squared of the regressions performed on the ETFs of the sample, the less the variation in the returns of said ETFs would be explained by the returns of its underlying Benchmark.

A consequence of this view of R-square is that it can proxy a tracking efficiency measure. Indeed, Figure 8 provides a visually clear example of why ($1-R^2$) or (1-Correlation) could be seen as easy-to-compute proxies for Tracking Error.

ETF	Obs.	Alpha	Beta	R^2	Std. Dev. (TE5)	t-value	Significance
1	1305	0.000014	1.005999	0.9847	0.00138	142.44	1%
2	934	0.000178	0.757552	0.5939	0.07830	16.03	1%
3	987	0.000045	0.862064	0.7382	0.00612	28.01	1%
4	1305	0.000027	1.019306	0.9939	0.00099	196.11	1%
5	1305	0.000012	1.011036	0.9912	0.00112	302.18	1%
6	987	0.000189	0.820681	0.8262	0.00642	21.24	1%
7	1305	0.000019	1.014635	0.9935	0.00102	223.59	1%
8	1305	0.000020	0.981704	0.9881	0.00161	121.25	1%
9	1305	0.000002	1.016040	0.9816	0.00152	120.11	1%
10	733	0.000077	0.899967	0.7259	0.00667	22.84	1%
11	733	0.000052	0.987815	0.7629	0.00716	25.05	1%
12	1305	0.000016	1.002608	0.9832	0.00159	191.6	1%
13	1305	0.000002	0.960715	0.9415	0.00251	27.25	1%
14	1305	0.000012	1.016656	0.9879	0.00123	257.11	1%
15	1305	0.000015	1.009727	0.9894	0.00134	190.85	1%
16	1305	-0.000011	0.965428	0.9857	0.00139	82.46	1%
17	1305	0.000007	1.009387	0.9785	0.00162	81.43	1%
18	1305	-0.000022	1.032061	0.7166	0.00786	33.3	1%
19	1305	0.000008	1.024547	0.9658	0.00212	100.26	1%
20	1305	0.000413	0.568643	0.4078	0.01099	12.14	1%
21	1305	0.000016	1.014123	0.9896	0.00128	241.85	1%
22	1305	-0.000074	0.919037	0.9458	0.00308	53.79	1%
23	1305	0.000019	1.006751	0.9828	0.00165	73.96	1%
24	1305	0.000017	1.001106	0.9725	0.00185	72.57	1%
25	1305	0.000029	1.010439	0.9929	0.00106	201.92	1%
26	1305	0.000018	0.989181	0.9840	0.00138	80.86	1%
27	1305	0.000012	1.012985	0.9828	0.00142	205.44	1%
28	1305	0.000014	1.016565	0.9895	0.00118	246.08	1%
29	1305	-0.000005	1.009787	0.9878	0.00122	248.11	1%
30	1305	0.000017	1.015184	0.9935	0.00097	375.53	1%

Table 8: Standard OLS regression of ETF returns and Benchmark returns

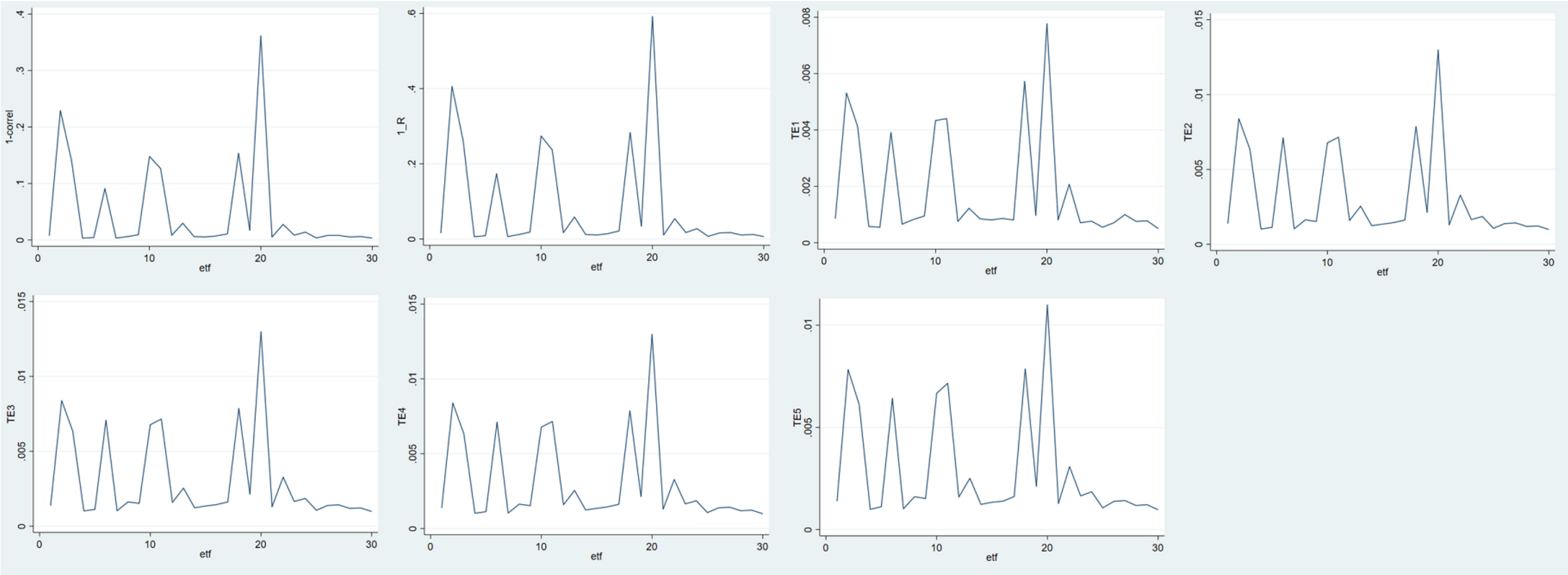


Figure 8: Graphical representations of TE1, TE2, TE3, TE4, TE5, ($1-R^2$) and (1-Correlation) values on the ETF sample

This computation of the Tracking Error through linear regressions also allows the analysis of the alphas produced following Charupat & Miu (2013), as a positive alpha (intercept) would mean that the ETF overperformed compared to its benchmark, while a negative alpha underperformed.

From Table 8, we can see that out of the 30 ETFs in the sample, 26 seem to have outperformed their benchmark, and only 4 underperformed. If we link this with the actual cumulative returns observed and displayed in Table 7, we see that 5 ETFs have underperformed, and not 4 as displayed by the intercept of the regressions. To reconcile both tables, we must look at the standard deviation of the returns of the 13th ETF. Indeed, the intercept alpha computed being very close to 0 at 0.000002; and the standard deviation being 0.0000771, the 95% confidence interval is [-0.0001497; 0.0001528], which means that the alpha could be negative and therefore reconcile both tables.

This overrepresentation of overperforming ETFs compared to their underlying benchmarks is linked to the choice of ETF providers to follow Net Returns benchmark, and our choice to use them as well in the analysis to provide an accurate representation of the actual returns earned by retail investors holding those types of ETFs and being subject to withholding taxes.

Both the alphas from the regressions and the actual cumulative returns displayed in table 7 allow us to rank the ETFs in the sample based on their overperformance. The Tracking errors computed in the previous sections also allow us to rank the ETFs in the sample based on their Tracking Efficiency. As reconciling both ranking is neither easy nor intuitive (we have shown that the most overperforming ETF in the sample was also the one with the highest Tracking Error), we must resort to other means to provide a simple measure of the actual risk-adjusted attractiveness of each of the ETFs of the sample that can be easier to compute and understand by non-institutional investors.

To do so, the Information ratio based on the Tracking Error of the 5-year daily returns of the ETFs of the sample was computed. The original Sharpe-ratio was deemed to have too many flaws in taking both upside and downside volatility into account (as detailed in Section 3.3) and was therefore not computed. The Sortino ratio was not computed either as it requires the investor's minimum acceptable return level to be computed but is of interest when it comes to ETF selection by retail investors. The Information ratios are computed based on Equation (11) and are displayed in Table 9.

ETF	Average daily TD	TE2	Information ratio
1	0.00164%	0.001380	1.19%
2	0.00074%	0.008392	0.09%
3	0.00084%	0.006338	0.13%
4	0.00339%	0.001020	3.33%
5	0.00158%	0.001128	1.40%
6	0.00088%	0.007104	0.12%
7	0.00240%	0.001031	2.32%
8	0.00111%	0.001630	0.68%
9	0.00085%	0.001527	0.56%
10	0.00258%	0.006774	0.38%
11	0.00463%	0.007154	0.65%
12	0.00169%	0.001587	1.06%
13	-0.00140%	0.002545	-0.55%
14	0.00180%	0.001240	1.45%
15	0.00202%	0.001347	1.50%
16	-0.00124%	0.001450	-0.85%
17	0.00100%	0.001620	0.61%
18	-0.00100%	0.007865	-0.13%
19	0.00177%	0.002138	0.83%
20	0.00481%	0.012977	0.37%
21	0.00211%	0.001295	1.63%
22	-0.01260%	0.003281	-3.84%
23	0.00216%	0.001650	1.31%
24	0.00179%	0.001854	0.97%
25	0.00325%	0.001065	3.05%
26	0.00142%	0.001381	1.03%
27	0.00167%	0.001428	1.17%
28	0.00213%	0.001192	1.79%
29	-0.00008%	0.001227	-0.07%
30	0.00226%	0.000985	2.29%

Table 9: Information ratios

Based on the Information Ratios (IR) computed on Table 9, we can see that the rankings of the ETFs in the sample change from those based on either only performance or Tracking ability. Indeed, the best performing ETF in term or returns (ETF 20) only has an IR of 0.37% because of its very high Tracking Error. The rankings based on performance, tracking ability and IR are summarized in Table 10.

Rank	Ranking Perf.	Ranking TE2	Ranking IR
1st	20	30	4
2nd	11	4	25
3rd	4	7	7
4th	25	25	30
5th	10	5	28
6th	7	28	21
7th	30	29	15
8th	23	14	14
9th	28	21	5
10th	21	15	23
11th	15	1	1
12th	14	26	27
13th	24	27	12
14th	19	16	26
15th	12	9	24
16th	27	12	19
17th	1	17	8
18th	5	8	11
19th	26	23	17
20th	8	24	9
21st	17	19	10
22nd	6	13	20
23rd	9	22	3
24th	3	3	6
25th	2	10	2
26th	29	6	29
27th	18	11	18
28th	16	18	13
29th	13	2	16
30th	22	20	22

Table 10: ETF Rankings

In using the Information ratio to assess the risk-adjusted attractiveness of the ETF in the sample, it becomes apparent that the allocation process is not a simple task for the average retail investor, and that merely looking at the returns of the ETF or its Tracking Error is not enough to properly assess the quality of an ETF.

7.4 Historical evolution of the Tracking Error

Section 7.1 to 7.3 have highlighted the apparent growth of the Tracking Error over the last couple of years for the ETFs of the sample, but only assessing it visually is not enough to properly assert this growth.

As a mean of testing this hypothesis that goes against our initial intuition, the average daily Tracking Difference and the average daily Tracking Error for the whole sample was computed using the three Tracking Error computation methods displayed in Equation (3) to (5). Once computed, the samples were divided in two periods of the same duration (652 observations). This results in eight sets of 652 observations, two sets of two and a half years for each of the four Tracking Errors computed.

To be able to use a paired t-test to check if the differences in Tracking Error are statistically significant, both samples must be tested for normality, as t-test have to be run on pseudo-normal data sets.

A Shapiro-Wilk test for normality was therefore performed on the eight sets of data, whose results are displayed in Table 11.

Variable	Obs.	W	V	Z	Prob>z
TE1_1	652	0.90523	40.522	9.004	0.00000
TE2_1	652	0.71835	120.424	11.653	0.00000
TE3_1	652	0.72196	118.884	11.621	0.00000
TE4_1	652	0.71748	120.796	11.660	0.00000
TE1_2	652	0.85163	63.440	10.094	0.00000
TE2_2	652	0.66345	143.900	12.086	0.00000
TE3_2	652	0.65467	147.653	12.149	0.00000
TE4_2	652	0.66307	144.063	12.089	0.00000

Table 11: Shapiro-Wilk test for normality

As the median value for V is 1 for normal population and as we have large values in the samples, this indicates nonnormality. The p-value therefore rejects the null hypothesis of normality with high confidence.

Non-normal distribution in the sample should prevent us from running the t-test needed to test the differences in Tracking Errors, but due to the central limit theorem, this does not affect the validity of the t-test as there is large amount of daily Tracking Error observations (652 observations for each Tracking Error computation method).

Figure 9 presents the distributions in the samples to visually assess their pseudo-normality.

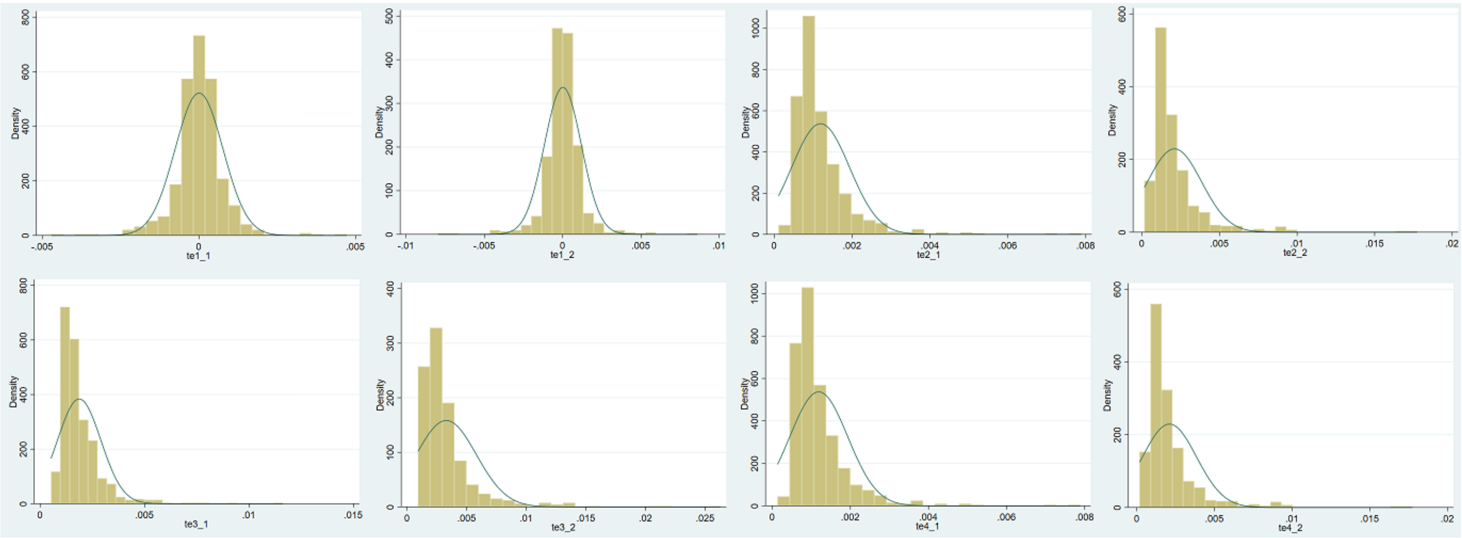


Figure 9: Distribution of the observations in the Tracking Error samples

A skewness and kurtosis test was also run to test for normality and displayed the same results as the Shapiro-Wilk one.

Four t-tests were therefore performed to test the equality of means of each of the sample combinations. The results of those t-tests are displayed in Table 12.

Variable	Difference	t	P-value	Significance
TE1_2 – TE1_1	5.39E-06	0.0968	0.9229	-
TE2_2 – TE2_1	0.0009117	11.9118	0.0000	1%
TE3_2 – TE3_1	0.001413	12.6717	0.0000	1%
TE4_2 – TE4_1	0.0009115	11.911	0.0000	1%

Table 12: t-tests results

Table 12 presents the results of the 4 paired t-tests performed to test the null hypothesis that there was no difference between the first 2.5 years of Tracking Errors reported and the following 2.5 years.

The t-test was found not significant for the measure of the Tracking Difference, but all three other t-test were found to be significant at the 1% level, and the mean difference for all three of the tests was found to be positive.

This positive mean of the difference supports our previous idea that Tracking Error had significantly increased over the studied period, which comes as a surprise compared to our initial hypothesis that Tracking Error would be reducing over time thanks to improvements in the models used by the ETF providers, the democratization of the Core-Satellite approach to ETF investing and the overall democratization of ETFs.

This can potentially be explained by the turmoil in the markets provoked by the Covid-19 and the general uneasiness characterizing the stock markets over the past couple of years because of the increasing environmental, geopolitical and economical tensions, or by other determinants of the Tracking Error.

To try to provide retail investors with some more information about the main determinants of the Tracking Error on the sample of Core ETFs, a panel data regression and OLS regressions will be performed in the following Section.

7.5 Determinants of the TE : Regression and results

In order to test the impact of the main determinants of the Tracking Error described in the relevant literature and detailed in Section 3.2, a regression analysis needs to be performed.

The first step in running a regression analysis is to define the variables needed, which has been done in section 5.3. Once those variables have been selected, the data set needs to be created to start working.

In our case, the variables needed to create the dataset are the following :

- The monthly Tracking errors calculated using daily returns for each of the four computation methods.
- A variable controlling for the size of the ETF, proxied here by the natural logarithm of the monthly AUM following the standards used in the relevant literature, whose original value is in million €.
- A dummy variable controlling for the replication method of the fund, it equals 1 is the fund follows a synthetic replication, 0 if it is physically replicated.
- A measure of the Volatility of the returns of the ETF, computed as the standard deviation of the fund's daily returns.
- The monthly dividend yields for the ETF.
- The Total expense ratio of the fund.
- A month variable to control for time.

From this list of variables, the monthly Tracking Errors were computed for 60 months for each ETF in the sample using the daily returns extracted from Bloomberg, the monthly AUM were extracted from Bloomberg using the FUND_TOTAL_ASSETS function, the dummy variable was generated on Stata using a 'gen(x) if y' syntax, the Volatility variable was computed using the daily returns extracted from Bloomberg, the TER of the funds were extracted from the prospectuses and checked using the FUND_EXPENSE_RATIO function in Bloomberg; the monthly dividend yields had to be manually extracted from the DVD function in Bloomberg as neither the DVD_HIST_ALL nor the DVD_SH_LAST functions were working properly on Bloomberg.

Before considering the different tests that need to be performed to find the most suitable regression model, the potential effects of collinearity between the independent variables must be checked. The correlation matrix of the independent variables was generated in Stata and is reported in Table 13. Correlation coefficients can be used as a way to find whether multicollinearity is an issue in a multivariate regression. Indeed, multicollinearity is believed to be an issue when two or more predictors have an absolute correlation coefficient above 0.6.

	SIZE	IS_SYNTH	TER	VOL	DY
SIZE	1.0000				
IS_SYNTH	0.0063	1.0000			
TER	0.2634	0.3870	1.0000		
VOL	-0.0266	-0.0202	-0.0413	1.0000	
DY	0.0239	-0.1251	-0.0414	-0.0141	1.0000

Table 13: Independent variables correlation matrix

As the highest correlation coefficient observed between the independent variables is only 0.3870, we can interpret the results of the regression without worrying about multicollinearity being an issue.

After taking care of this multicollinearity issue, several tests can be run to try to determine the most suitable regression to be performed with our data.

A F-test was first performed after declaring the data to be panel data. The F-test tests the hypothesis that all coefficients are equal to zero. As the F-value for the test is equal to 79.84, we have a p-value inferior to 0.0000 that reaches statistical significance, which means that the null hypothesis that all coefficients are equal to zero is rejected and that a panel-effect exists in our dataset.

A Hausman test comparing the use of Fixed Effects regression to Pooled OLS resulted in a chi2 of 407.22 and a p-value of 0.0000, therefore reaching significance and rejecting the null hypothesis that the preferred model is Pooled OLS.

A Lagrange Multiplier test was then used to test the null hypothesis of a zero variance of the fixed effects and resulted in a chibar2 of 13 048.42 and a P-value of 0.0000, therefore reaching significance and rejecting the null hypothesis.

Finally, another Hausman test comparing the use of Fixed effects to Random Effects resulted in a chi2 of 21.92 and a p-value of 0.0005, therefore reaching significance and rejecting the null hypothesis that the preferred model is Random Effects.

As a consequence of those tests, the preferred model to run the regression on our dataset is a Panel Data Fixed Effects model.

Several aspects of our model are nonetheless problematic regarding the implementation of our regression using a panel data fixed effects model. Indeed, the TER variable is time-invariant as it only changes between ETFs but never over time within the same ETF. The SYNTH variable controlling for the replication method is both a dummy variable and a time-invariant one, as it only changes between ETFs but never within an ETF.

Both time-invariant and dummy variables get omitted when running a Fixed Effect Panel Data regression, because as they do not change over time, the related coefficient will not be estimated under the fixed effects specification. This issue can be solved through several means, most notably using the Mundlak approach or the

Generalized method of moments (GMM) in Stata. Because of this added complexity given the specificities of our dataset and my limited knowledge of those regression techniques, it was decided to resort to simple OLS to study the determinants of the Tracking Errors computed on our sample.

The results of the 4 OLS regressions ran on the model of Equation (13) are displayed in Table 14.

	TE1	TE2	TE3	TE4
SIZE	0.0002903***	0.0003277***	0.000325***	0.0003304***
	(0.000099)	(0.00013)	(0.00013)	(0.0001302)
SYNTH	0.0015586***	0.0020362***	0.002042***	0.0019878***
	(0.0001848)	(0.0002424)	(0.0002421)	(0.0002441)
TER	-0.8365704***	-1.105273***	-1.103014***	-1.092531***
	(0.0726162)	(0.0991219)	(0.099019)	(0.0991404)
VOL	0.1760451***	0.2516977***	0.2517282***	0.248692***
	(0.281309)	(0.0363689)	(0.036371)	(0.0366508)
Div_Y	-0.0030188	-0.0023858	-0.0022762	-0.0023805
	(0.0066669)	(0.0088613)	(0.0088496)	(0.0088167)
# of Obs.	1703	1703	1703	1703
# of funds	30	30	30	30
R^2	0.2278	0.2426	0.2425	0.2364

Notes : standard error in parenthesis. Significance expressed by stars (*= 90%; ** = 95%, ***= 99%). The constant was not reported.

Table 14: OLS regressions for TE1, TE2, TE3, TE4

The OLS regression of the first Tracking Error (TE1) as dependent variable has an R^2 of 0.2278. This can be interpreted as : 22.78% of the variation in the Tracking Error can be explained by the independent variables used in this model.

All of the independent variables are significant at the 1% level, except for the Dividend Yield that is not significant at any relevant level.

The SIZE variable is significant at the 1% level and has a positive coefficient, which is surprising as most of the existing literature reports a negative coefficient for this variable. Indeed, they explain this negative coefficient as the result of economies of scale, reduced transaction costs and optimization for larger funds. A potential explanation for this positive coefficient, that can be understood as 'for a 1 unit increase in size, the TE1 will increase by 0.0002903 units', could be some kind of slackness appearing in the largest ETFs, now considering themselves big enough to not have to be as efficient as others. In any case, this coefficient being relatively small, this result tends to support the findings of Rowley & James (2015) that doubted the significance of size and for whom size only had a marginal effect on the Tracking Error.

The SYNTH variable is significant at the 1% level and has a positive coefficient, which supports the observations made by Drenovak & Al (2014) that states that a Physical ETF, and especially a Full replication one is supposed to have the smallest Tracking Error, even smaller than for synthetic ETFs.

The TER variable is significant at the 1% level and has a negative coefficient, which is very surprising given the nature of the management fees that directly affect the performance of the fund and causes them to underperform by the same amount. The results of our regression can be interpreted as 'for a 1 unit increase in TER, the TE will decrease by 0.8365704 units'. This surprising result could be explained by several factors. First, given the increased competition in the ETF space, it might be possible that funds that try to track their underlying benchmark as precisely as possible are subject to more transaction costs due to more frequent rebalancing of the securities. Second, the increased attractiveness for ETF in the current investment landscape might have led some providers known for the excellence of the Tracking ability of their funds to raise their TER so as to be compensated for this skill, which might also explain this negative coefficient.

The VOL variable is significant at the 1% level and has a positive coefficient. This positive coefficient supports the findings of Qadav & Yagil (2012) that found the daily volatility of the ETF to be positively correlated with the Tracking Error. This makes sense as the impossibility for the ETF, ceteri paribus, to perfectly follow its index is amplified when dealing with high volatility.

The Div_Y variable is not significant at any relevant level and will therefore not be interpreted. This non-significance is probably linked to the small amount of ETFs in our sample that actually distributed dividends over the studied period. Most of the ETFs in our sample were Accumulating ones, and the few Distributing ones were mostly distributing dividends only semi-annually, which led to a very small sample of Dividends to use in the regressions.

The other regressions performed with TE2, TE3 and TE4 led to very similar results, all the independent variables being significant at the same level, all of the coefficients having the same sign and being in the same order of magnitude, it seems unnecessary to reinterpret them a second time.

The R^2 of the regressions with TE4, TE3 and TE2 are nonetheless slightly higher than the one of the regression with TE1, them being at 23.64%; 24.25% and 24.26% respectively.

7.6 Best performers and ex-ante recommendations to investors

Based on the different analyses performed in this thesis, a ranking of the ETFs in our sample can be made and several piece of advice can be shared.

First of all, based on an analysis of both the overperformance and the Tracking ability of the ETFs, the following ranking was made for our sample, based on their risk-adjusted performance. Column three nonetheless provides the TER of each of the funds to finalize the optimal decision-making process.

Rank	Ranking IR	TER
1st	4	0.07%
2nd	25	0.05%
3rd	7	0.20%
4th	30	0.18%
5th	28	0.11%
6th	21	0.18%
7th	15	0.25%
8th	14	0.15%
9th	5	0.12%
10th	23	0.18%
11th	1	0.07%
12th	27	0.11%
13th	12	0.10%
14th	26	0.10%
15th	24	0.20%
16th	19	0.20%
17th	8	0.35%
18th	11	0.05%
19th	17	0.15%
20th	9	0.25%
21st	10	0.05%
22nd	20	0.15%
23rd	3	0.04%
24th	6	0.04%
25th	2	0.12%
26th	29	0.25%
27th	18	0.20%
28th	13	0.30%
29th	16	0.25%
30th	22	0.25%

It is worth noting that the two best ETFs in terms of risk-adjusted performance are also among those with the lowest TER, making them the absolute best risk-adjusted options in our sample (based on past information not necessarily translating into the future).

Several guidance can be derived from this research to try to put some light on how to avoid more easily the pitfalls of dominated products. When it comes to ETFs, the following tips can be applied based on the research done in this thesis :

- Differentiate ETFs following Gross, Net or Price return Indexes, as their performance and seeming tracking ability can greatly vary based on the Index type.
- Rather than focusing solely on either overperformance or Tracking ability, easy-to-compute risk-adjusted measures can be used to have a clearer view of the best performing ETFs, such as the Information ratio or the Sortino ratio.
- Some easily accessible measures such as correlation or R^2 can to some extent be used as a proxy for ETF tracking efficiency if other more accurate measurements are not available.
- ETFs Tracking ability is significantly influenced by market turmoil, and market crashes greatly increase the Tracking Error of ETFs, particularly those with the lowest correlation to the movements of their benchmark.
- Care should be taken when looking at confusing prospectuses that are often overcomplicated and often use information that can appear misleading to non-institutional investors.
- When selecting ETFs, pay attention to the following determinants affecting their Tracking ability : TER, Replication Method (focus on Physical replication), Volatility (focus on low volatile indexes), and somewhat but less importantly Size (AUM).

8 Conclusion

Through the various analyses performed in this thesis, the main goal was to study the performance and the Tracking ability of a sample of Europe Core ETFs so as to provide useful advice to help mitigate the development of asset-allocation by retail investors into dominated products.

8.1 Summary of findings

After providing the reader with an overview of ETFs, from their conception and their recent development to their main performance measures, a review of the recent literature covering ETFs, Tracking Error, and other performance measures was made and a sample of ETFs was created to be studied.

By computing the Tracking Error using four different measures for each ETF of the sample and by computing the Average Tracking Error for the whole sample, the study showed the existence of a significative Tracking Error for European Core ETFs, mostly positive for this sample due to the distinction made between Net, Gross and Price return indexes but still relatively low when compared to specialized ETFs. The study also showed significant deviations between ETFs in terms of Tracking ability.

Computing those Tracking Errors also allowed for the analysis of big deviations from the benchmarks occurring during market crashes and amplifying the low tracking ability of certain ETFs.

With the computed average Tracking Errors, dividing our sample into two equal time periods of 2.5 years allowed us to run a paired sample t-test to check whether the Tracking error had stayed the same or not, and we found that contrary to our initial beliefs, the average daily tracking error had increased over the last period.

Computing the Tracking error as the standard deviation of the OLS regression with ETF returns as dependent variable and Benchmark returns as independent variable allowed us to verify that the intercept $\propto$ from said regression was a good estimator of whether an ETF outperformed its benchmark or not, and we found that correlation and R^2 could be useful approximate proxies for the Tracking Error in case of lack of more precise measurements.

Studying the performance of the ETFs in our sample using the Information Ratio as a risk-adjusted measure allowed us to assess the overall quality of an ETF in an easier and more accurate way, and we provided a ranking of the ETFs of the sample based on this measure.

Furthermore, the study built on the existing literature concerning the determinants of the Tracking Error by running OLS regressions and found Size, TER, Volatility and Replication Method to be explanatory variables, but failed to demonstrate the same for Dividend Yields. The coefficient for the Size and TER variables were also found to be negative, contrary to the majority of the literature, even though some other previous studies found the same negative coefficient.

8.2 Limitations

As it was mainly aimed at European retail investors, the scope of this study focused solely on European Core ETFs while most of the literature used in this study focuses on U.S, Asian or world indexes. As such, there are significant differences in results between the different geographical areas, and specificities of the European markets might have been forgotten as not considered in the previous literature.

Another potential limitation of this study is the choice to use Net Return Indexes instead of Gross or Price Returns. Indeed, this choice might be responsible for some of the variations in result between this study and previous ones, as Net Return Indexes tend to lead to bigger excess returns for the studied ETFs, and consequently to potential biases in the study.

Finally, the timeline chosen for the study encompassed the recent Covid-19 financial crash, which might have biased the calculations of our Tracking Errors and Tracking Errors Averages over this period of high instability.

8.3 Potential research opportunities

Some of the topics scratched during this study might be of interest for further research. The impact of Covid-19 on Tracking Errors should be further explored to see how ETFs reacted during a financial crisis of this scope.

The apparent growth of the Tracking Error would also deserve to be further studied, as this goes against the idea of improved tracking ability due to increasing competition between ETF providers and other technical improvements.

Finally, the subject of using ‘quick-and-easy’ measures such as the correlation or the R^2 as proxies for the Tracking Error would deserve more attention as well, as it could help retail investors make informed decisions without getting lost in the large amount of information available on ETFs.

Appendices

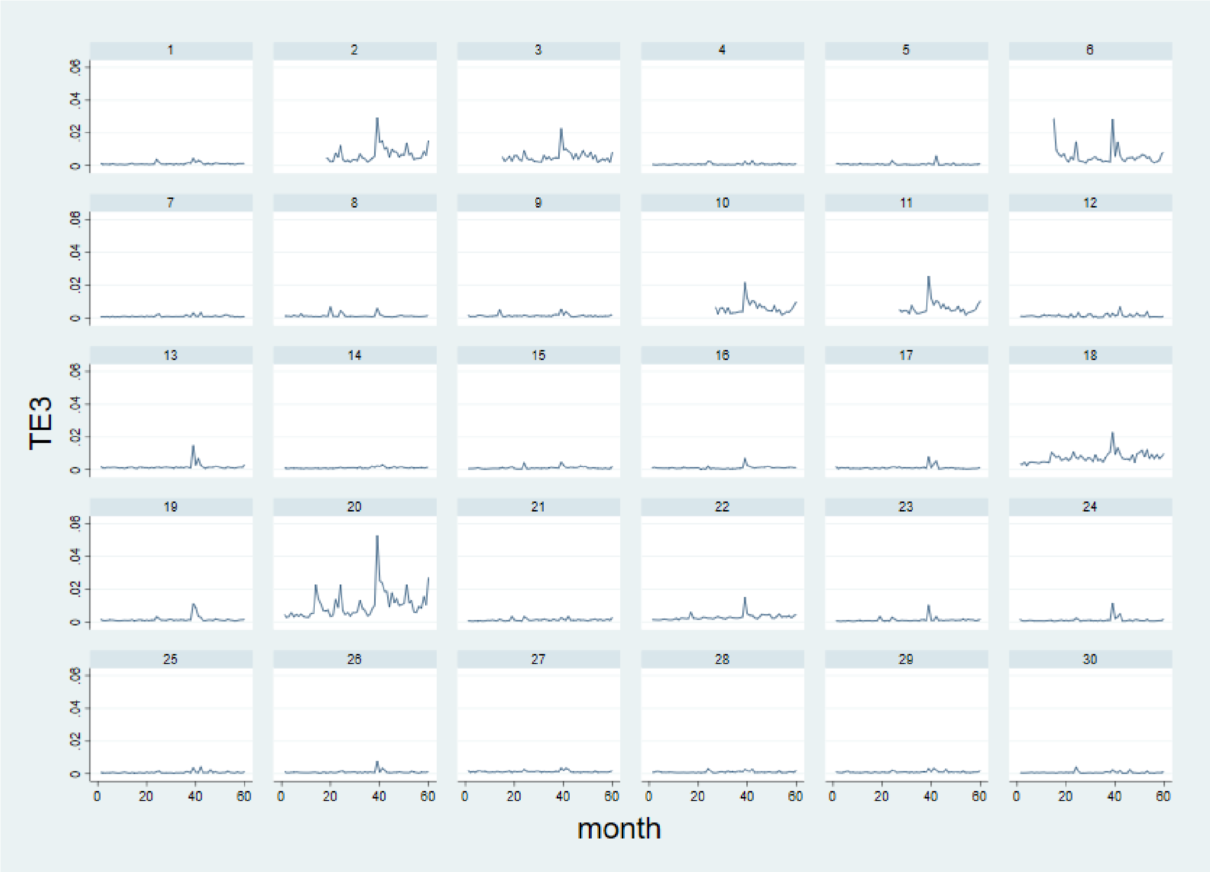
Appendix 1: Table and Graphic representation of the difference in returns between an ETF, its Net Return Benchmark and its Gross Return Benchmark. The data was extracted from Bloomberg using the COMP command, and the returns calculation is set on the NAV. The difference in returns presented runs from December 30th, 2017, to December 31st 2021, a 5-year period.

Range 12/30/2016 - 12/31/2021 Period Daily No. of Period 1827 Day(s) Table

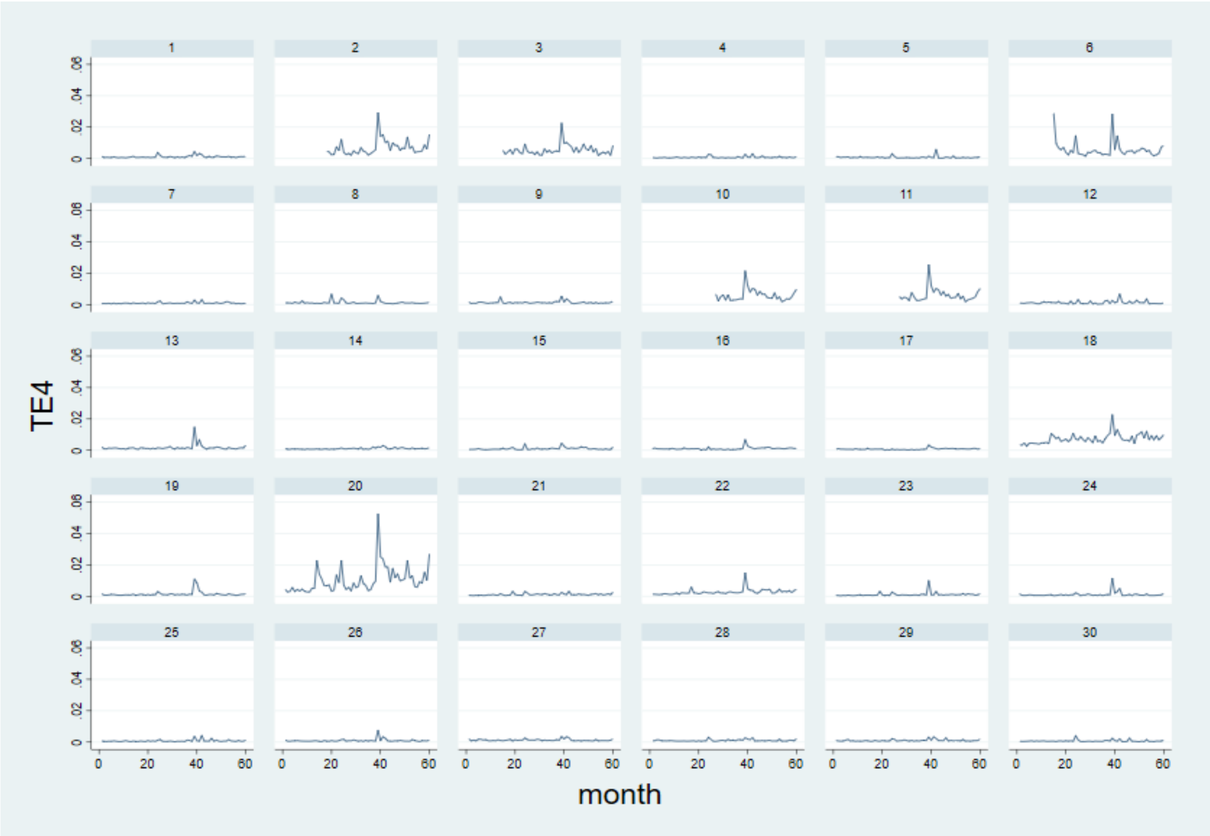
	Security	Currency	Price Change	Total Return	Difference	Annual Eq
1)	MEUD FP Equity	EUR	55.04%	55.04%	--	9.16%
2)	SXXR Index	EUR	53.19%	53.19%	-1.85%	8.89%
3)	SXXGR Index	EUR	57.57%	57.57%	2.53%	9.51%
4)						
5)						
6)						

1M 3M 6M YTD 1Y 2Y 3Y 5Y 10Y

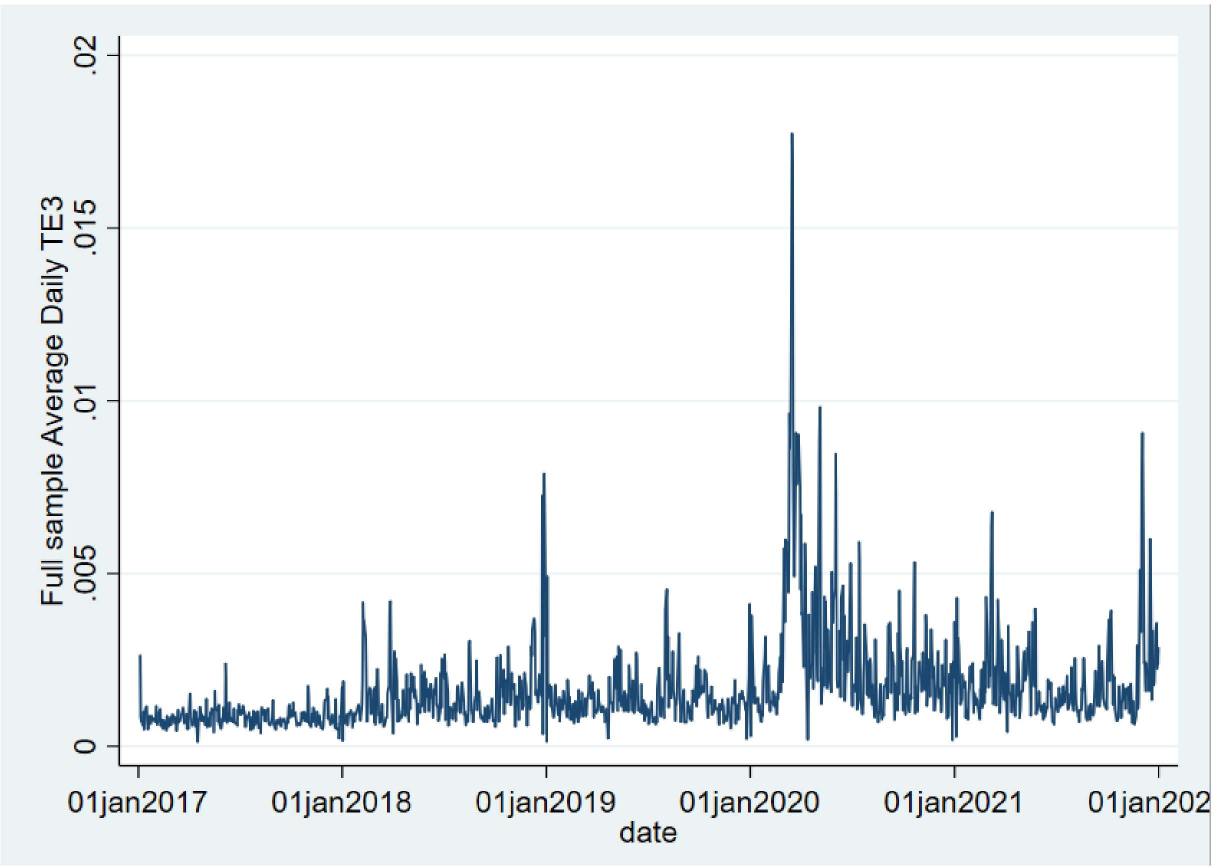
Appendix 2: Tracking Error (TE3) over time of all ETFs in the sample.



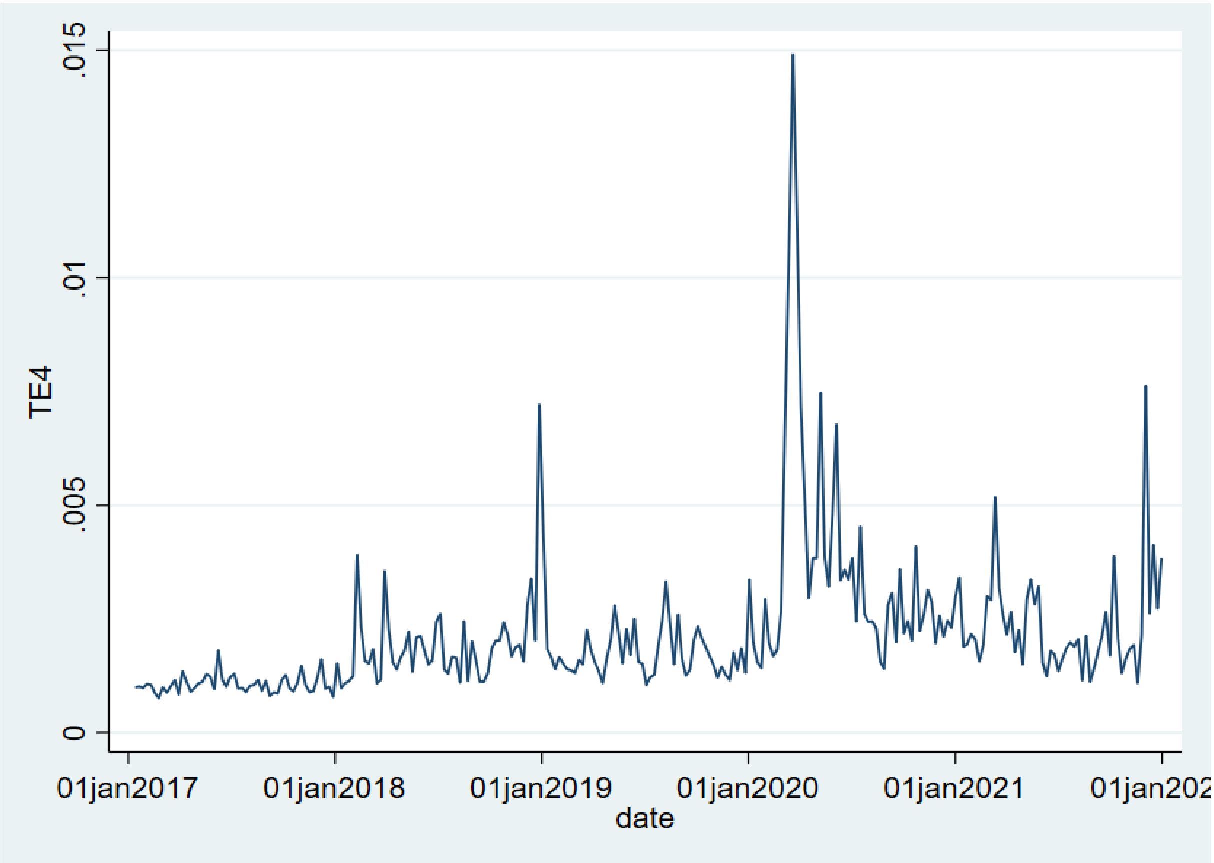
Appendix 3: Tracking Error (TE4) over time of all ETFs in the sample.



Appendix 4: Sample Average Daily Tracking Error 3



Appendix 5: Sample Average Weekly Tracking Error 4



Bibliography

Agapova, A., 2011. Conventional mutual index funds versus exchange-traded funds. Journal of Financial Markets, 14(2), 323-343.

Amenc., Goltz., 2007. ETFs in Core-Satellite Portfolio Management. ETFs and Indexing Sep 2007, (1) 75-88.

Amenc., Goltz., Grigoriu., 2010. Risk Control through Dynamic Core-Satellite Portfolios of ETFs: Applications to Absolute Return Funds and Tactical Asset Allocation. EDHEC-Risk Institute

Bajpai, P., 2021. What Are Synthetic ETFs ? Nasdaq, Inc

Ben-David., Itzhak., Franzoni, F., Moussawi, R., 2017. Exchange-traded funds, Annual Review of Financial Economics 9, 169–189.

Ben-David., Itzhak., Franzoni, F., Moussawi, R., 2021. Competition for Attention in the ETF Space. NBER Working Paper No. 28369. January 2021, Revised January 2022. JEL No. G12,G14,G15

Bordalo, Pedro, Gennaioli, N., Shleifer, A., 2016. Competition for attention, Review of Economic Studies 83, 481–513.

Buetow, G. W., Henderson, B. J., 2012. An empirical analysis of exchange-traded funds. Journal of Portfolio Management, 38(4), 112.

Charupat N., Miu P., 2013. Recent developments in exchange-traded fund literature. Managerial Finance, vol. 39, no. 5, pp. 427–443.

Chaudhry, A., Johnson, H. L., 2008. The efficacy of the Sortino ratio and other benchmarked performance measures under skewed return distributions. Australian Journal of Management, 32(3), 485-502.

Chu., P.K.-K., 2013. Tracking errors and their determinants: Evidence from Hong Kong exchange traded funds. University of Macau, Macao.

Chu., P.K.-K., 2016. Analysis and Forecast of Tracking Performance of Hong Kong Exchange- Traded Funds: Evidence from Tracker Fund and X iShares A50. Review of Pacific Basin Financial Markets & Policies 19, 1–26.

Cresson, J.E., Mike Cudd, R., Lipscomb, T.J., 2002. The early attraction of S&P index funds: is perfect tracking performance an illusion? Managerial Finance 28, 1–8

Brown D.C., Cederburg, S., Towner, M., 2021. Dominated ETFs

Drenovak, M., Urošević, B., Jelic, R., 2014. European Bond ETFs: Tracking Errors and the Sovereign Debt Crisis: European Bond ETFs. European Financial Management 20, 958–994

Elia, M., 2012. Tracking Error of Traditional and Synthetic European Exchange-Traded Funds. Working Paper, University of Turin, Turin.

Elton, E.J., Gruber, M.J., Comer, G., Li, K., 2002. Spiders: Where Are the Bugs? The Journal of Business 75, 453–472.

Frino, A., Gallagher, D.R., 2001. Tracking S&P 500 Index Funds. The Journal of Portfolio Management 28, 44–55

Frino, A., Gallagher, D.R., Neubert, A.S. and Oetomo, T.N., 2004. Index design and implications for index tracking, Journal of Portfolio Management, Vol. 30 No. 2, pp. 89-95.

Flood, C., 2021. “ETF assets close to $10tn after second year of record growth”, Financial Times

Gastineau., Gary, L., 2004. The Benchmark Index ETF Performance Problem. The Journal of Portfolio Management, 30(2), 96–103.

Gennaioli, Nicola, Shleifer, A., Vishny, R., 2012. Neglected risks, financial innovation, and financial fragility, Journal of Financial Economics 104, 452–468.

Goodwin, T.H., 1998. The Information Ratio. Financial Analysts Journal, 54(4), 34–43.

Hill, J. M., Nadig, D., Hougan, M., 2016. A Comprehensive Guide to Exchange-Traded Funds (ETFs)(a summary). Research Foundation Publications, 2016(1), 18-23.

Johnson, W. F., 2009. Tracking errors of exchange traded funds. Journal of Asset Management, 10(4), 253-262.

Johnson, B., Bioy, H., Kellett, A., Davidson, L., 2013. On the Right Track: Measuring Tracking Efficiency in ETFs. The Journal of Index Investing 4, 35–41

Johnson, S., 2021. "Historic trend reverses as ETF fees head higher", Financial Times

Larsen, G., and Resnick, B., 1998. Empirical Insights on Indexing, Journal of Portfolio Management, 25, pp. 51-60.

Lee., Charles, M.C., Shleifer, A., Thaler, R.H., 1991. Investor sentiment and the closed-end fund puzzle, Journal of Finance 46, 75–109.

Methling, F., Von Nitzsch, R., 2018. Thematic Portfolio Optimization: Challenging the Core Satellite Approach, Journal of Global Optimization

Novick, Barbara., 2017. How index funds democratize investing, Wall Street Journal (January, 8).

Osterhoff, F., Kaserer, C., 2016. Determinants of tracking error in German ETFs–the role of market liquidity. Managerial Finance, 42(5), 417-437.

Pope, P.F., Pradeep, P.K., 1994. Discovering errors in tracking error. Journal of Portfolio Management 20, 27.

Qadan, M., Yagil, J., 2012. On the dynamics of tracking indices by exchange traded funds in the presence of high volatility, Managerial Finance, Vol. 38 Iss 9 pp. 804-832

Rompotis, G. G., 2006. Evaluating the Performance and the Trading Characteristics of iShares. Working Paper, University of Athens, Department of Economics, Athens.

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

Rompotis, G. G., 2012. The German Exchange Traded Funds. The IUP Journal of Applied Finance, 18(4), 62-82.

Rowley, J.J., Kwon, D.T., 2015. The Ins and Outs of Index Tracking. The Journal of Portfolio Management 41, 35–45.

Saunders, Kent T., 2018. Analysis of International ETF Tracking Error in Country-Specific Funds. Atlantic Economic Journal

Sin,R., Psarofagis, A., 2021. Stigma surrounding synthetic ETFs should be put to rest for good. Bloomberg Intelligence

Tang, H., Xu, X. E., 2014. Tracking Performance of Leveraged and Regular Fixed-Income ETFs. The Journal of Fixed Income, 23(3), 64.

Vanguard Group, Inc., 2021. Vanguard’s guide to Core-Satellite investing

iShares Core - Investor Guide: https://www.blackrock.com/us/financial-professionals/literature/investor-guide/ishares-core-essentials-guide-en-us.pdf