INSTITUTO POLITÉCNICO DE LISBOA
INSTITUTO SUPERIOR DE
CONTABILIDADE E ADMINISTRAÇÃO DE
LISBOA

EXCHANGE TRADED FUNDS STUNNING PERFORMANCE: WHAT DRIVES ABNORMAL RETURN

Hugo Hilário Varela
Master in Financial Analysis

Lisboa, 31 de outubro de 2018

INSTITUTO POLITÉCNICO DE LISBOA
INSTITUTO SUPERIOR DE CONTABILIDADE E ADMINISTRAÇÃO DE LISBOA

EXCHANGE TRADED FUNDS STUNNING PERFORMANCE: WHAT DRIVES ABNORMAL RETURN

Hugo Hilário Varela

Dissertação submetida ao Instituto Superior de Contabilidade e Administração de Lisboa para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Análise Financeira, realizada sob a orientação científica de Carlos Pinheiro.

Constituição do Júri:

Presidente Doutor Joaquim Ferrão

Vogal Doutor Carlos Pinheiro

Arguente Doutora Sónia Bentes

Lisboa, 11 de Junho de 2019

Acknowledgments

First, I would like to thank my thesis supervisor, Carlos Manuel Pinheiro, for guiding and supporting me over these last two years.

The experiences, ideas, and opinions that you share with me through this process have been extremely valuable.

You are an example of excellence as a researcher, mentor, and role model for life.

I'm also in debt with my friends, colleagues, co-workers, professors (João Fernandes, Filipe Cizeron, Sérgio Ferreira, Francisco Barreto, Érica Casanova …) and all the people that in somehow helped me through this process and collaborate for the development and conclusion of this research project.

I would also like to thank my family for the love, support, and constant encouragement provided throughout this entire process, from the very beginning to the end.

Resumo

Este trabalho de investigação visa analisar o desempenho dos Exchange-Traded Funds que replicam os principais índices bolsistas Europeus e Americanos. Testamos a capacidade dos ETFs de bater o mercado através da análise do parâmetro de alfa de Jensen e averiguamos o impacto da liquidez no retorno dos ETFs, aplicando o modelo de vários fatores de Fama e French.

No geral os ETFs que replicam os principais índices bolsistas Americanos exibem um desempenho superior aos ETFs Europeus. Contudo a maioria dos ETFs analisados não conseguem bater o mercado. As diferenças no retorno parecem estar ligadas às disparidades nos níveis de liquidez existentes nos 2 mercados. Os resultados empíricos demonstram que o subgrupo de ETFs que replicam índices bolsitas Europeus exibem maior sensibilidade às oscilações nos níveis de liquidez do que o subgrupo dos ETFs Americanos.

Os resultados obtidos são consistentes com modelos convencionais de análise de risco-retorno e também revelam o poder explicativo das variáveis de liquidez no retorno dos ETFs, que até então não foi investigado, bem como os fatores de dimensão e valor desenvolvidos em Fama e French (1993).

Palavras Chaves: Exchange-Traded Funds performance; Abnormal Return; Fama-French; Liquidity; Capital Assets Pricing Models (CAPM); Reward-To-Volatility ratio.

Abstract

This research paper investigates the performance of Exchange-Traded Funds (ETFs) linked to major U.S. and European stock indexes. We test their abilities to generate an abnormal return, and we employ a modify Fama-French three-factor model to examine the role played by ETFs' specific characteristics such as liquidity in determining ETFs returns.

In terms of overall performance, we found that ETFs that replicate the main stock indexes in the U.S. outperforms the European ones. However, the majority of ETFs analyzed were unable to beat the market. Differences in the expected return rate might be explained by the disparities in the liquidity levels between the two markets. The results show that the subsample of European ETFs is more sensitive to changes in liquidity than the U.S. subsample.

The results are consistent with conventional risk-return models, and it reveals the high predictability power of liquidity variables in determining ETFs returns, which so far was left uninvestigated by the literature.

Keywords: Exchange-Traded Funds performance; Abnormal Return; Fama-French; Liquidity; Capital Assets Pricing Models (CAPM); Liquidity; Reward-To-Volatility ratio.

Table of contents

Index of Figures

Table Index

List of Acronyms and Abbreviations

ETFs – Exchange-Traded Funds;

CAPM – Capital Asset Pricing Model;

Bp – Basis point;

S&P500 – Standard and Poor's 500;

SPDR – Standard & Poor's Depository Receipt;

U.S. – United States;

AP – Authorized Participant

NAV – Net Asset Value

GEM – Global Emerging Markets

NYSE – New York Stock Exchange

DJIA – Dow Jones Industrial Average

1. Introduction

Increasingly more and more investors and asset managers worldwide are investing larger volumes on alternative financial products combining outstanding performance, lower cost structure, ease in trading and low exposure to idiosyncratic risk.

Thus, tradeable funds, such as Exchange-Traded Funds (henceforth ETFs), are becoming very popular in the securities market, due to their ability to generate substantial returns at lower cost and to provide investors with prompt diversification and exposure to a large number of market indexes.

ETFs represent a basket of securities that can be traded as common stock or any other tradable security in the stock exchange. It is managed on a daily-basis for tracking purposes so that they can replicate market indexes.

Another distinctive feature of ETFs is related to their trading characteristics, in what their exposure to price volatility on a daily-basis is concerned. On top of this, bearing in mind that they can easily be bought or sold in the market, ETFs exhibit a higher daily liquidity than most traditional mutual funds.

Despite the high popularity of ETFs nowadays, studies concerning the risk-return structure of European ETFs and its relationship with microstructure factors, such as liquidity are still scarce. The existing literature is more focused on the performance of ETFs designed to mimic major stock market indexes in the U.S.

The limited evidence on the performance of European ETFs, as well as the absence of empirical research, constitutes two major motivational factors for our study.

European ETFs is by far less liquid than U.S. counterpart and most of its shares is negotiated in over-the-counter markets, by institutional investors. Contrarily, U.S. retail investors contribute significantly to the funds flows toward U.S. ETFs, using transparent and more liquid on-exchange trading channels to sell or buy ETF shares.

Low scale retail investors playing in European and U.S. market face different liquidity constraints and therefore they should be examined separately. The demand for ETFs in Europe is growing at a very faster pace, but it still lags U.S. ETFs. Stafford (2016) argued that the major differences between European and U.S. ETFs have to do with the fact that European ETFs are rarely used by retail investors and its often negotiated under off-exchange contracts. Assets under management and liquidity disparities between them are very significant, he documented that in 2016 assets held by European ETFs peaked at $567bn, while U.S. ETFs recorded $2.4tn during the same period, as we can see in **Figure 1.1**, and in **Figure 1.2**.

<Figure 1.1 Liquidity in the U.S. ETF Market>

<Figure 1.2 Liquidity in the European ETF Market>

The importance of microstructure factors in determining assets returns, such as stocks liquidity is well present in the literature and it continues to trigger the attention of many academicals who pursue new sources of risk and more efficient asset valuation models, other than the market risk provided by the Capital Asset Pricing Model (henceforth denoted CAPM).

Amihud (2002) documented a positive and significant impact of illiquidity on the expected stock returns, and more recently Ibbotson, Chen, Kim, and Hu (2013) demonstrated that liquidity is viable and not "easily beaten" alternative to other investment styles (i.e. size, value/growth, and momentum).

We contribute to the existing literature in two ways, by first providing an in-depth analysis on the performance of the ETFs tracking major European and American benchmark indexes, thus fulfilling the existent gap in the literature focused mainly on the performance of ETFs designed to track major stock market indexes in U.S., while the performance and pricing efficiency of European ETFs are poorly mentioned in the literature.

Second, we employ cross-sectional regression models to provide empirical evidence of the main driven forces of ETFs expected returns rate, which so far, to the best of our knowledge no study has specifically investigated the main determinants of ETFs returns.

We found that our European subsample of ETFs is more sensitive to shocks in liquidity than their American counterparts. The estimations reveal a strong and robust effect of liquidity on the expected excess returns of European equity ETFs, meanwhile, U.S. subsample of ETFs exhibit a lower intensity in response to oscillations in spreads.

Particularly for the European subsample of ETFs, we found robust and convincing results. The estimated coefficients of spread and Amihud illiquidity factor are both statistically and economically significant, while for the U.S. subsample of ETFs no reliable coefficients of spread, and Amihud illiquidity factor was found, regardless the regression model used.

The expected return rate of European ETFs, as well as their ability to outperform the market, are clearly undermined by sudden changes in liquidity, as they are more prone to incorporate overall changes in bid-ask spread, exhibiting a highly sensitive response function to shocks in liquidity and inferior performance compared to the ones tracking major benchmark indexes in the U.S.

2. Literature review

2.1 General Framework of ETFs

Exchange-Traded Funds are a relatively recent phenomenon. It all started in 1993, with the successful launch of SPDR (Standard & Poor's Depository Receipt), a unit investment trust holding portfolio poised to match and track the performance of S&P 500 (Standard and Poor's 500). SPDR's incredible success hails from its special and revolutionary trading characteristics – investors can trade SPDR throughout the day, unlike more traditional mutual funds that can only be traded at the end of the day when their net asset value is calculated.

The systematic growth of investment funds and tradeable "pools of money", that allow investors to trade individual shares in the exchange market without incurring in high costs and expensive managerial fees imposed by actively managed funds, will pave the way for more sophisticated and highly developed passive managed funds.

During the past decade, as investors started to realize the advantages of ETFs and its tradability features, the appetite for ETFs peaked, and both the volume of assets under management (AUM) and the number of ETF players reached record-breaking numbers. Thereby, it was only a matter of time until ETFs took over the markets, increasingly conquering new investors, disseminating brand new products into it, and providing a diversified and well-structure range of products to investors, thus fulfilling on a consistent basis their appetite for highly profitable, diversifiable and global range financial products.

Twenty years ago, the assets under management of the main Sponsors were almost negligible, whereas nowadays they exceed $3tn. In the last year, U.S. mutual funds experienced a massive outflow of $130.7bn according to the annual report provided by research firm Morningstar (see, Authers and Newlands (2016)).

There is no doubt, that ETFs triggered systematic shifting in the investment vehicles used by investors, from traditional actively managed open-ended mutual funds, towards passive managed ETFs, as shown in **Figure 2.1**.

<*Figure 2.1 ETFs Take Over*>

In turn, the number of ETFs products traded worldwide, including equity, commodities, and fixed income ETFs, reached huge proportions. In 2000 there were approximately less than 300 ETFs products available worldwide; while in 2016, the number of ETF products was surprisingly close to 5,000 (see, **Figure 2.2**).

<*Figure 2.2 Number of ETFs Products Available Worldwide*>

Based on the information above, analyzing the performance of ETFs is undoubtedly a topical issue as it can contribute to the financial literature, providing a better understanding of the forces at play in terms of ETF's return, volatility, efficiency, and liquidity. ETFs brought about dramatic changes and have had a significant impact on stock trading and capital markets worldwide.

2.1.1 Creation and Redemption process of ETFs

Another important feature of ETFs is their ability to be created or redeemed to meet and satisfy large-scale investors' needs. In a summarized way, the creation and redemption mechanisms of ETFs reflect how ETFs gain exposure to the market, which confer ETFs with special agility and efficient responsiveness mechanisms to deal properly with sudden changes in the market. The Creation and Redemption mechanism allows ETFs to be more flexible and capable to provide additional liquidity and less expensive transactions to investors.

Through the creation mechanism, an Authorized Participant (AP) can acquire the securities that the ETF intends to hold. The acquired securities are delivered to the ETF provider, and in turn, the AP demands from the providers a block of equally valued ETF shares, called a creation unit. Thus, whenever an ETF provider wants to disseminate and create shares available for traders or launch new products to meet investors demand, he can resource to a market maker with strong buying power (AP). The AP will acquire the required securities in exchange, in most cases in blocks of 50,000 ETFs shares, which can be resold whenever it seems profitable to do so.

The AP can remove ETF shares from the market by purchasing a significant amount of ETF shares to form a creation unit, and then deliverer those shares to the ETF provider, receiving in turn, the same value converted into underlying securities of the fund. This is called the redemption mechanism. This feature is very important because it can force the readjustment process of ETF’s shares price, whenever there is an imbalance regarding the market value of ETF and the value of its underlying securities. For example, if many investors want to buy an ETF, triggering a significant shift in demand toward ETF, the ETF shares prices might rise at higher speed and scale comparatively with its underlying securities. Thus, in this situation the AP is expected to intervene forcing the overpriced ETF shares prices to come down to their fair value, by purchasing the underlying shares that compose the ETF, selling afterward the shares in the open market pushing the price down.

When the market value of an ETF is above their net asset value (NAV) or the value of each share of its underlying securities we can say that, the ETF is trading at a premium. In turn, when ETF shares are trading below the price of its NAV, the ETF is trading at discounts. Thus, when the first situation occurs, an AP must trigger the redemption process, removing ETF shares from the market, preventing market asymmetries that might arise, in case investors decide to close immediately their long positions on ETF holdings, dumping them into the market.

2.2 Past Studies

The research on previous studies follows three mains strands: (i) performance and tracking efficiency models, (ii) liquidity as an investment style, and (iii) conventional asset pricing models.

2.2.1 Performance and Tracking Efficiency

A large part of the debate over traditional index funds and ETFs concerns their ability to beat the market, performing relatively better than the inherent benchmark, without compromising their major goal (tracking the index as closely as possible at lower cost).

However, this ability and operational proficiency do not apply to all ETFs. Gastineau (2004) documents that ETFs linked to popular U.S. stocks indexes, such as Russel 2000 and S&P 500 exhibited a weak performance as compared to their competitors. Gastineau (2004)'s results show that iShares Russel ETF underperformed the inherent benchmark by 53 basis points (bps) or 0,53%, in 2001, while the Vanguard Small Cap Index Fund beat the Russel 2000 benchmark by 76bps (0.76%) in the same period. To explain the weak performance of ETFs as compared to competing products, like mutual funds, Gastineau (2004) focused on the management policy undertaken by ETFs' portfolio managers, which he believed to be the source of a structural problem. Whenever there is a change in the composition of the inherent index, the portfolio manager will have to decide when and how to adjust the fund formation in or order to reflect the changes occurred in the index.

This process of matching the weights of the securities held by the ETF accordingly to the size of the index is supported by the creation and redemption mechanisms. As stated by Gastineau (2004), the transaction cost incurred each time a market manager or Authorized Participant (AP) decides to create new shares or extinguish existing ones, cannot be ignored. In fact, the price charged to investors tends to reflect the administrative costs related to handling the creation or redemption basket, as the ETF manager attempts to replicate the returns and the risk of the underlying market index, as closely as possible.

Elton et al. (2005) examined the performance of SPDR (Standard & Poor's Depository Receipt), commonly known as a spider, a famous and popular ETF that tracks the S&P 500 index. They found that, on average, the NAV (Net Asset Value) return on the Spider underperforms the S&P 500 return, by 28bps (0.28%) per year. Elton et al. (2005) posit that the difference in performance hails from the cost structure and the tracking efficiency of spiders.

The management fee, as well as other management expenses, constitute a clear disadvantage for ETF investors. The return of a spider is reduced in a significant way by the transaction costs incurred in replicating the index, adjusting the weights of their holdings every time there is a change in the composition of the index. Once again, the intrinsic structure of ETF, involving the creation and redemption of shares seems to affect their ability to outperform the underlying benchmark.

The replication strategy should be such that, at any point in time the number of stocks held by the ETF and the stocks in the S&P 500 feature an exact match in terms of weights. This situation does not always occur, and when it does not, the efficiency of the ETF in replicating the index is clearly undermined. In the meantime, Elton et al. (2005) found no significant evidence that connects the tracking error to the inferior performance of the Spiders relative to the benchmark.

Frino and Gallagher (2001) report the struggle faced by index-oriented fund managers on a daily basis, as they attempt to replicate the underlying index without incurring in substantial transaction cost. They highlight the fact that tracking error in performance or inaccurate replication strategy are unavoidable due to market frictions. Theoretically, the task of index managers to realigning the weights of securities in the portfolio, trading index securities whenever there is a change in the composition of the underlying index, does not have to be a “difficult job”, as the market itself converges to a state of equilibrium.

Unfortunately, the real-world market conditions prevent, in many ways, transactions to take place smoothly. First, there is a cost involved, every time the index manager decides to buy or sell new shares. Second, transactions or trading schemes involving a considerable number of shares of stocks do not occur instantaneously, but rather depend on the investors’ willingness to perform quickly such trade and on the size of the stocks entering and exiting the index.

Therefore, market timing to incorporate a certain transaction is a really important issue for index managers. According to them, the magnitude and the variation of the tracking error arise directly from the volatility of the securities that constitute an index. Dividend distribution policies and the timing to effectively pay to investors can also cause tracking error.

Examining the 42 S&P 500 index mutual fund, Frino and Gallagher (2001) document the presence of tracking error in performance. Their results show that the cross-sectional average absolute tracking error[1] is equal to 5.9bps (0.059%) per month before expenses.

The tracking error estimated as the volatility of the difference in returns is approximately 8.0bps (0.08%) per month or 27.6bps (0.276%) per annum. The correlation between dividend payments to the holders of S&P 500 securities and the tracking error is positive and statistically significant. There are signs of seasonality in the tracking error. The estimated tracking error is significantly higher in the months of January and May. By comparing active funds with passive S&P 500 index mutual fund, on average the latter outperform active funds, exhibiting a risk-adjusted excess return after expenses higher than that of large-cap active mutual funds.

Svetina and Wahal (2008) used a sample of 584 ETFs to demonstrate that on average ETFs underperform their underlying benchmark, due to the transaction costs. Domestic equity ETFs representing approximately 60% of the sample examined underperform their benchmark by 0.26 percent per year, while international equity ETFs underperform by 0,19 percent per year. Contrary to Elton et al. (2005)'s study, Svetina and Wahal (2008) found that retail index-oriented mutual funds are outperformed by ETFs at about 0.30 percent per year, on average. Concerning the accuracy in tracking the underlying index, the results are consistent with the literature. For the domestic equity ETFs and international equity ETFs, they estimate a tracking error of 0.47% and 1.13% respectively.

Shin and Soydemir (2010) found similar results, using three different methods to calculate the tracking error of 26 ETFs. They found evidence of persistence in tracking errors, which explains why the examined ETFs underperform their benchmark. This study highlights the inability of the passive investment strategy to beat the market.

Kostovetsky (2003) provides a consistent comparative analysis between ETFs and conventional index funds, by developing a simple one-period model which proved to be very useful in examining substantial differences between the two investment vehicles.

[1] The absolute tracking error accounts for the difference in returns between the index portfolio and the underlying benchmark. Another common metric of tracking error also used in this paper is the standard deviation of the difference in returns between the index portfolio and the underlying index.

The competitive cost advantage of ETFs over indices relates to their ability to disseminate and eliminate shares in the market through a subtle and highly effective mechanism (creation and redemption), which allows large investors to purchase considerable amounts of ETF shares without any exposure to direct liquidity costs in the secondary market, offering only, in turn, a portfolio of securities matching the target index in weights and value.

Apart from the fund transaction costs and exploitation of arbitrage opportunities, which are quite irrelevant for ETFs due to the above mentioned motives, Kostovetsky (2003) points out three other relevant costs that can probably influence investors' choice between ETF or index funds. Rebalancing costs due to changes in the index or corporate activity (i.e. merging or acquisition operation, forcing the merged or acquired company to leave the index) are relevant and affect both, ETFs and index funds, in the same proportion. Shareholder transaction costs are a major issue for ETF investors who must pay a commission to the broker in order to purchase shares in the secondary market (with exception of large investors). Contrarily, most index funds charge no commissions on transactions (i.e. no-load index funds). Finally, the tax efficiency is the third factor that affects ETFs and index funds in a different way. While tax efficiency associated with capital gains favors ETFs, index fund investors are negatively affected by the tax burdens. The in-kind redemption process of ETFs allows for the washing out of the most appreciated stocks, thus capital gains distribution is a rare operation for most ETFs.

Poterba and Shoven (2002) compare the pre-tax return and after-tax return on the SPDR trust (largest ETF that tracks the S&P500), with the returns on the largest equity index fund, the Vanguard Index 500 fund. The results suggest a superior pre-tax and after-tax performance of the Vanguard Index 500 as compared to the SPDR trust. The superior performance of the index fund might be related with profitable trading strategies adopted by the fund managers (i.e. purchasing shares of companies that intend to make part of the S&P 500 at the announcement date).

With the same goal of comparing the performance and the tracking error of an ETF with an index mutual fund, Bello (2012) investigated small caps ETFs and index mutual funds that track the same benchmark (Russel 2000 index). His results suggest that small-cap ETFs are less diversified than index mutual funds, because ETFs generally overinvest in the top-ten companies they hold, thus only a small percentage of their investments is allocated into “less-valuable” securities. As expected, ETFs have lower expense ratios and high portfolio turnover than index mutual funds.

To evaluate ETFs and index mutual funds’ relative performance, Bello (2012) uses Jensen’s alpha and Sharpe information ratio. Consistent with previous research he found that, despite the superior performance of index mutual funds relative to small caps ETFs, both underperformed their benchmark. Concerning the tracking efficiency, Bello (2012) found that ETFs exhibit a larger tracking error than index mutual.

Despite the similarities and discrepancies between these two instruments (conventional index mutual funds and ETFs), their coexistence is inevitable. They both provide to investors a diversified portfolio and alternative investment opportunities, competing against each other in a very sustainable way, exploiting systematically effective means to reduce managerial and transaction costs. Agapova (2011)’s study shows that conventional funds and ETFs are substitutes, but not perfect substitutes for one another. Which means that despite the innovative marketability and tradability features brought about by ETFs, conventional funds are still not easily replaceable. The competition and coexistence between these two products are sustainable and useful for investors, who benefit from increasing competition pushing prices down.

Cremers, Ferreira, Matos, and Starks (2016) hypothesize that the competitive pressure brought by index funds not only forces the active funds to be more cost attractive but also to be more risk efficient, generating positive alpha. They found that on average active funds charges lower (higher) fees in markets with more explicit indexing funds (closet indexing[2]). Therefore, the entry of new index funds conceived explicitly to track the performance of a benchmark index, does benefit investors, as the fee charged by active funds tend to decrease. Consistent with the underlying hypothesis of the paper, Cremers, Ferreira, Matos, and Starks (2016) found that the average alpha generated by skilled active funds’ managers is higher in countries in which low-cost passive managed funds are more popular, and lower in markets dominated by closet indexing.

Pricing deviations and situations of imbalance between ETFs’ price and the market value of its underlying securities are another topical issues. Thirumalai (2003) finds

[2] Closet indexing is a terminology used to describe funds that falsely claim to actively purchase investments but use passive strategy instead. They normally charge high fees, as if they are active, and yield average returns similar to a benchmark index.

evidence of price deviation from NAV, by analyzing the pricing efficiency of passive and active ETFs. DeFusco, Ivanov, and Karels (2011) focused primarily on the mismatch between the price of ETF and the price of the market index being tracked. They used the Engle and Sarkar (2006)'s approach to calculate the price deviation as the difference between the price of the market index and the price of ETF. They found that price deviation of the three most liquid ETFs (Spider, Diamonds, and Cubes) are predictable and stationary. The results suggest that, on average, the spider ETF is priced at 29 cents above S&P 500's price, Diamond's price is 8 cents higher than the price of Dow Jones Industrial Average (DJIA), and the Cubes ETF is priced at 25 cents below its underlying benchmark. Once again dividend accumulation and distribution policy, as well as the intrinsic nature of ETF and their price formation mechanisms, are proclaimed to constitute the main source of the problem, causing pricing deviation and inefficient managerial performance.

Ivanov (2013) extends the previous work of DeFusco, Ivanov, and Karels (2011), using high-frequency data. Ivanov (2013 found empirical evidence of high-frequency pricing deviations between ETFs and their underlying market indexes. Price deviations of DIA, SPY, and QQQ amounts to 0.0429, -0.0743 and 0.4298 respectively. These findings cast doubts on the efficiency of highly sophisticated monitoring computers and algorithmic-based trading programs that are responsible for clearing any signs of mispricing of ETFs. Pricing inefficiency of ETFs can benefit investors, who can make a riskless profit using arbitrage strategy (i.e. take a short (long) on the overpriced (underpriced) ETF and buy (sell) the cheapest (expensive) basket of securities). In the meantime, arbitrage opportunities and situations of discrepancies in prices are quickly overcome, as the market converges to equilibrium.

Hilliard (2014) found that the median long-term premium of U.S. equity ETFs tends to converge to zero, which confirms the high efficiency of ETFs' arbitrage pricing mechanism.

Thus, for some ETFs the creation and redemption process are being executed in an efficient way, meaning that profitable arbitrage opportunities (when the NAV of the ETF differs from its shares prices) are very limited due to the effectiveness of the creation and redemption strategy. International ETFs face a more complex process. In some countries, there are barriers to transferring and transacting shares overseas, making the arbitrage mechanism for international ETFs costlier and riskier than for domestic ones.

Engle and Sarkar (2006) highlight the problem faced by international ETFs in preventing their shares to be traded above or below NAV. Their findings indicate persistent and frequent premiums or discounts related to international ETFs.

Research on the performance of ETFs linked to European, Asian, Australian or other emerging markets is scarce. The absence of empirical research and the limited evidence on the performance of European and Asian ETFs are a clear limiting factor for our study. However, we should recognize that even today, innovative and dynamic products like ETFs are still new concepts for most investors.

Gallagher and Segara (2006) investigate the performance and trading characteristics of classical index-oriented ETFs in Australia. They provide a consistent analysis regarding the ability of ETFs to replicate persistently the Australian Stock Exchange, by comparing their tracking error volatility with others index funds operating in the market. Gallagher and Segara (2006) use two classical methods of calculating the tracking error, widely addressed in the literature (absolute difference in returns and standard deviation of the difference in returns between the ETF and the underlying index), and they found empirical evidence of significant tracking error across all ETFs and index-oriented funds tracking the Australian Stock Exchange index. In contrast with the previous studies, they found no evidence of superior or inferior performance of ETFs relative to their underlying benchmark.

Further results show that the wholesale index funds examined exhibit higher tracking error than ETFs, and the variation between the NAV and the quoted ETFs' prices are very small and do not occur frequently. Blitz and Huij (2012) investigate the overall performance of ETFs designed to track global emerging markets (henceforth denoted GEM) equity indexes (i.e. China, Brazil, India, South Africa and Russia). They argued that tracking errors of those ETFs are much more likely to be higher than the tracking errors of the ETFs that mimic developing markets. The tendency of relative high tracking errors for GEM ETFs was explained based on the following arguments. First, the dispersion in stocks returns is larger for emerging markets; and second, the trading costs arising from low levels of stocks' liquidity is much higher in those emerging markets than in developed markets. The results confirm their hypothesis, as they found that GEM ETFs exhibit substantially higher levels of tracking errors than their developed markets counterparts. Consistent with the existing literature mainly focused in the investigation of ETFs linked to U.S. equity indexes, Blitz and Huij (2012) found that the GEM ETFs fall short of their benchmark indexes by around 85bps (0.85%), due to expected drag on returns caused by expenses and dividends taxation.

Regarding ETFs that provide passive exposure to European indices, the lack of empirical evidence is even more evident. Performance and pricing efficiency of European ETFs are poorly mentioned in the literature, apart from Blitz, Huij, and Swinkels (2012)'s research focused on the performance of European ETFs and index funds, not much has been said. Blitz, Huij, and Swinkels (2012) posit that dividend withholding taxes as the main determinant of European ETFs and index funds inferior performance relative to their benchmarks. They concluded that dividend taxes are at least as important as fund expenses ratios in explaining European ETFs and index funds relative weak performance. They found that, on average, expense ratios drag down passive European funds' performance by 56bps (0.56%), while dividend taxes contribute to funds' performance at -48bps (-0.48%) per year.

Milonas and Rompotis (2006)'s study provides empirical evidence of the performance and marketability of 36 ETFs tracking the Swiss Stock Exchange index, an important player in the European ETF market. They applied a single index model to analyze Swiss listed ETFs ability to beat their underlying benchmark. The estimated coefficients are in line with previous studies – no evidence of superior performance was found – the mean alpha coefficient of the entire ETF sample was negative and statistically significant. The mean beta coefficient was significant and below one, meaning that the replication strategy adopted by those ETFs failed to fully mimic the return and volatility of their corresponding indexes, which can partially explain why they, on average, underperformed their benchmark by 7bps (0.07%). By regressing ETFs average daily return on management fee ratio, using cross-sectional regression analysis, Milonas and Rompotis (2006) found that the performance of ETFs is significantly and negatively affect by management fees. The impact of management fees on ETFs tracking errors was found to be positive and statistically significant.

Yiannaki, (2015) investigates the performance and tracking accuracy of ETFs domiciled in Ireland and Luxembourg. Consistent with previous studies, he found no evidence of abnormal returns or superior performance. The estimated Jensen's alpha parameters were insignificant and close to zero, as well as the tracking errors. They concluded that despite their successful start, Ireland and Luxembourg listed ETFs underperform their benchmark.

Before we address the next topic of our literature review, we would like to close this discussion over ETFs performance and tracking efficiency, summarizing the findings and its implications in our research.

As spotlighted in literature, ETFs' ability to replicate perfectly the performance of their inherent benchmark, matching its return and volatility, is not always accomplished, thus in line with the existing literature, we can say that tracking error is unavoidable.

The effectiveness of the creation and redemption process confers ETFs a trivial competitive advantage over conventional index funds, allowing them to be less expensive and more tax efficient.

These mechanisms also prevent ETFs shares to be traded above or below their underlying NAV, thus profitable arbitrage opportunities are very limited. However, findings concerning ETFs' pricing efficiency are ambiguous, there is not a consensus yet.

Unlike European ETFs, U.S. listed ETFs were targeted by many academicals who documented their inferior performance, relative to their inherent benchmark. The magnitude of their underperformance was commonly attributable to management expenses, rebalance and taxation costs.

In our study we will examine the overall performance and tracking efficiency of European and U.S. ETFs, using performance measures commonly used in the literature, such as, average return, risk-adjusted return, volatility metrics, tracking error (absolute and standardized versions), and we hope to find results that are consistent with previous findings for U.S. ETFs, and then compare it to European ETFs.

2.2.2 Liquidity

The role played by liquidity in defining the trading strategy adopted by investors is a matter of high importance among all asset classes. ETFs are no exception, indeed the number of investors seeking liquid and fast exposure to an entire market index is increasing sharply.

A recent study made by Greenwich Associates in collaboration with BlackRock (Greenwich Associates 2018), revealed that liquidity and volume represent one of the most important criteria in selecting ETFs among others available investment vehicles, according to the 80% of institutional investors surveyed.

Amihud and Mendelson (1986) introduce the bid-ask spread as a natural measure of liquidity and test its effect on asset pricing. They provided evidence of an increasing and concave relation between spread and asset returns. Which confirms their hypothesis – less liquid stocks outperform more liquid stocks, on average. The presence of clientele effect was also tested, Amihud and Mendelson (1986) demonstrated that investors can mitigate the effect of transaction costs picking stocks that best fit their expected holding periods or investment horizons. Investors and assets were ranked according to their expected holding periods and bid-ask spreads respectively. The estimated trading costs due to bid-ask spread was lower for investors who trade less frequently. Thus, the reduction in the expected return rate caused by illiquidity costs is lower the longer the security is held.

Fund managers and stock investors tend to adjust and customize their portfolio of assets to fit their investment horizons and liquidity goals. For active traders, the price discount due to illiquidity cost or future trading costs can represent a major problem and therefore they should invest in liquid securities. On the other hand, long-horizons investors can benefit from lower price discount, selecting more illiquid securities.

Therefore the clientele effect and its implications in capital markets documented by Amihud and Mendelson (1986) can guide investors trough portfolio selection processes, so they can achieve better results by incorporating market-microstructure factors, such as expected trading costs, into their investment strategy.

Datar, Naik, and Radcliffe (1998) use a different approach to measure liquidity[3] and found similar results, confirming the significant role played by liquidity in explaining stocks return behavior. Datar, Naik, and Radcliffe (1998) found a strong and significant negative impact of the turnover rate on stocks return. In line with previous study they show that illiquid stocks provide higher average returns.

[3] Datar, Naik, and Radcliffe, (1998) estimate a proxy for liquidity based on the stock turnover. The turnover rate of each stock was calculated by dividing the average trading volume in each month to the number of outstanding shares.

Amihud (2002) examines the impact of market illiquidity on stock excess return, based on a different measure of illiquidity, highly recommended in later research. The illiquidity measure was calculated as the average daily ratio of absolute stock return to dollar volume. He documented a positive and significant impact of illiquidity on the expected stock returns. For small stock firms, the estimated impact was stronger. Amihud (2002)'s results suggest that the risk premium required by investors as compensation for the risk is positively related with expected illiquidity. Which confirms his hypothesis that the expected excess return on a stock or the stock risk premium does mirror compensation for expected market illiquidity.

Ibbotson, Chen, Kim, and Hu (2013) posit that a well-established measure of liquidity such as stock turnover is a feasible indicator of U.S. stock returns. They found that the low-liquidity portfolio yields on average, a higher return than the high-liquidity portfolio, regardless of the size, value, and momentum of the sorted portfolio. Ibbotson, Chen, Kim, and Hu (2013) built a dollar-neutral liquidity model to demonstrated that liquidity is a viable and not "easily beaten" alternative to other investment styles, such as size, value/growth, and momentum. The estimated coefficients of their model suggest that the liquidity factor exhibit a stronger impact on returns than other styles factors. The estimated alphas were also positive and significant.

Idzorek, Xiong, and Ibbotson (2012) extend the framework of liquidity as an investment style, firstly tested at the security level, to mutual funds level. Idzorek, Xiong, and Ibbotson (2012) found that mutual funds composed mainly of less liquid stocks outperformed those that held more liquid stocks, on an average basis.

Chordia, Roll, and Subrahmanyam (2000) investigate commonality in liquidity and the key elements responsible for intertemporal changes in liquidity. Uncommon from previous research, focused primarily on the marketability components of individual assets, their research intends to analyze correlated movements in liquidity considering both individual and broad market trading activity. Chordia, Roll, and Subrahmanyam (2000) found empirical evidence of covariation between individual stock liquidity and market industry liquidity. They demonstrated that inventory risk and information asymmetry affect individual stock liquidity significantly.

Acharya and Pedersen (2005) investigate the impact of unexpected changes in liquidity on security's required return rate. They developed a liquidity version of the capital asset pricing model, that captures not only the conventional CAPM's expected return-beta relationship but also captures movements in prices caused by unexpected changes in the market and individual asset liquidity. Acharya and Pedersen (2005)'s liquidity-adjusted capital asset pricing model shows that investors' willingness to invest in a stock depends positively on the stock ability to provide higher returns when market liquidity crashes. Therefore, investors are willing to pay a premium for stocks that generate a higher return when the market illiquidity is greater.

So far, the observed literature is highly connected with our investigation purposes, and it provides strong theoretical and analytical support to our study. The importance of microstructure factors in determining assets returns, such as stocks liquidity is well present in the literature and it continues to trigger the attention of many academicals who pursue new sources of risk and more efficient asset valuation models, other than the market risk provided by the CAPM model. However, research focused on investigating the role played by liquidity in determining ETFs returns is scarce. The effect of liquidity on ETFs returns has not yet been tested, at least in a straightforward way.

Hegde and McDermott (2004) link improvements in the market liquidity to the introduction of index-tracking ETFs. Hegde and McDermott (2004) demonstrated that the market liquidity of the underlying DJIA (Dow Jones Industrial Average) 30 index stocks improve substantially over the first 50 trading days after the introduction of the DIA ETF, commonly known as DIAMONDS. They posit that information asymmetry and trading costs as the main causes. Contrarily De Winne, Gresse, and Platten (2014) found that the liquidity improvement observed after the introduction of ETFs are not driven merely by changes in the adverse selection costs but rather by changes in the order processing and order imbalance costs. Ivanov (2016) finds similar results, by showing that order processing cost represents the primary source of ETFs' bid-ask spreads.

There is no doubt that the increasing popularity of ETFs is largely attributed to the ease with which investors can obtain all the benefits of a diversified portfolio, without incurring high transaction costs. Broman (2016) relates high liquid ETFs to excess comovement in returns and demonstrates how attractive liquidity characteristics can facilitate excess returns. Broman (2016) hypothesize that short-term investors with correlated demand for investment styles are highly attracted by financial instruments like ETFs that combine both, easy access to all types of investment styles and lower transaction costs. He documented that similar styles ETFs exhibit positive commonality in misvaluation, while different styles ETFs comove negatively in terms of misvaluation.

Lee, Tseng, and Yang (2014)'s research provides evidence of commonality in liquidity for country ETFs. They found that the magnitude of commonality in liquidity varies with liquidity distribution, and it tends to be stronger during periods of financial crisis.

Marshall, Nguyen, and Visaltanachoti (2018) confirm the efficiency of liquidity proxies in capturing changes in effective and quoted spread, using a sample of 600 ETFs. Despite the differences between ETFs and stocks liquidity, Marshall, Nguyen, and Visaltanachoti (2018) demonstrated that liquidity proxies commonly applied on stocks are actually doing a good job in determining ETFs liquidity. The results suggest that effective spread on the Dow Jones Industrial Average ETF (DIA) is lower than the price-weighted effective spread of the stocks that comprise the underlying market index (DJIA).

Subrahmanyam (1991)'s seminal research show that a basket of securities exhibits on average higher liquidity than the underlying stocks. Subrahmanyam (1991) postulate that the transaction costs due to information asymmetry are lower in the market of index traded securities, such an ETF, than for individual security markets. He argued that liquidity traders are motivated by the need for immediate execution, which entails a cost. Intuitively liquidity traders prefer to invest in a basket of securities that provides lower information asymmetry and consequently lower transaction costs, rather than invest in a security individually.

Innovative financial instruments, such as an ETF, brought about substantial improvements in the market liquidity and information dissemination among traders. Liquidity and active traders, in general, are more likely to invest in low-cost index-tracking funds due to simple reasons, higher liquidity than the underlying stocks comprising the index and cheaper diversification instrument. Greenwich Associates (2018)'s survey shows that liquidity and exposure needs are the factors that matter the most when selecting ETFs (see **Figure 2.3**).

<Figure 2.3 ETFs Selection Criteria>

Low scale retail investors playing in European and U.S. market face different liquidity constraints and therefore they should be examined separately. The demand for ETF in Europe is growing at a very faster pace, but it still lags behind the demand for U.S. ETF. Meziani (2016) reported that the North America ETF market represents approximately 73% of the total ETF market, with $1,506,864 million in assets. While the Europe ETF market represents only 2% of the total ETF market, which corresponds to only $50,244 million of the total $2,061,095 million ETF market. Stafford (2016) show that the major differences between European and U.S. ETFs have to do with the fact that European ETFs are rarely used by retail investors and its often negotiated under off-exchange contracts.

Assets under management and liquidity disparities between the two markets are very significant, Stafford (2016) document that in 2016 assets held by European ETFs peaked at $567bn, while U.S. ETFs recorded $2.4tn during the same period.

European ETFs is by far less liquid than U.S. counterpart and most of its shares is negotiated in over-the-counter markets, by institutional investors. Contrary in the U.S., retail investors contribute significantly to the funds flows toward ETFs, and they use transparent and more liquid trading channels to sell or buy ETF shares. Recently published Flood (2018)'s newspaper article provides interesting market insights over European ETFs. Flood, (2018) reports that only 10 to 15 percent of the total assets invested in European ETFs came out of retail investors' pocket. Moreover, the number of passive shares classes available for sale to retail investors is very limited in most European countries. Exceptionally in the UK and Germany, the number of passive shares classes open to retail investors are very expressive, exceeding 8000. Equity and bond funds traded in the UK are less expensive than their European counterparts, the average fee for equity funds sold in the UK amounts to 66bps (0.66%) while the average fee in Europe is close to 189bps (1.89%).

The cost advantage of the UK funds combined with their vast range of passively managed products are pushing the demand up, attracting increasingly more retail investors into the funds market. Flood, (2018) document that financial advisers in most countries in Europe, with exception to UK and Netherlands, are less likely to recommend ETFs to their clients, while in the U.S. the reverse happens - ETFs are highly recommend by financial advisers.

For this paper, considering our investigation purposes and empirical hypothesis constructed based on the existing differences between U.S. and European funds markets, we decided to focus primarily on the potential relationship between liquidity measures and ETFs returns.

Therefore, we hypothesize that the easiness to which an investor can buy or sell ETFs can potentially affect its expected return rate. Our second model will test empirically the sensitiveness of most popular European and the U.S. listed ETFs' returns to changes in liquidity.

In the following section, we will start our discussion over return models. First, we will address the Classical CAPM tenets, and then we will go through factor models. Well-defined and practical concepts in financial markets, such as Investment opportunity set, passive strategies, incremental return per incremental risk, reward-to-volatility ratio, Jensen's alpha, risk-free asset, and capital market line, may emerge from this preliminary review over equilibrium pricing models.

2.2.3 Risk-Return Models

Even before starting to think about investment pay off, investors should be aware of their money allocation. So, they start off with a simple question – "How much should I own in stocks (risky assets) and how much should I invest in a risk-free asset or other types of safe money securities?" By answering this question, the investor automatically derives a simple and very useful finding in portfolio theory, the investment opportunity set or all the efficient investments choices available to an investor at a given period.

Sharpe (1964) shows that the investment plan chosen by an investor is efficient, if and only if, it can provide the highest expected return for a given level of risk and the lowest risk for the same level of expected return. To maximize their utility as a function of risk and expected return, investors can choose among all possible portfolios that lie on the investment opportunity curve.

This simple illustration of how investors can allocate their capital is a clear demonstration of a simple asset allocation choice, which can help us understand the dynamic of Index models and guide us through the derivation process of the Capital Market Line, according to the CAPM (see Lintner (1965)). The equilibrium conditions of the CAPM are restricted to a set of assumptions concerning investors' preferences and external market factors. CAPM assumes that all investors are rational, as they systematically apply the mean-variance theory developed in Markowitz (1952)'s seminal research[4], for optimization and utility maximization purposes.

Moreover, investors are considered to have identical investment horizons and share the same economic insight or homogeneous view of the market. So, they perform securities scanning using the same expected returns and covariance matrix of securities returns to create the optimal risky portfolio on the efficient frontier.

Investment decisions and portfolio construction strategies are defined according to the mean-variance theory. Therefore, the risk-averse investor will rationally invest in the efficient portfolio in order to obtain the lowest volatility for any given level of expected return. Income taxes and transaction costs are assumed to be zero (frictionless market) and there is no restriction on credit - short selling is allowed.

Finally, single and isolated investors' trade decisions do not affect the market's overall equilibrium and security prices. Thus, investors are assumed to be price-takers and the market itself embeds a set of conditions underpinning what is denoted as perfect competition.

With no restrictions on borrowing or lending, investors can increase their investment opportunities by investing some money in Treasury bills (borrowing or lending at the risk-free rate) and the remaining money in stocks.

Black (1972) shows that the efficiency of the CAPM can be substantially improved under circumstances where investors can neither have access to the riskless asset nor borrowing or lending at the risk-free rate. The model obtained by dropping one of the most important assumptions of the CAPM is consistent with the equilibrium models developed by Black, Jensen, and Scholes (1972) and contradicts Sharpe (1964) and Lintner (1965). The former contended that any point along the capital market line is desirable and attainable by investors if they behave rationally, adopting a diversified portfolio strategy.

[4] Harry Markowitz is the mind behind one of the most important pillars of modern portfolio management, the Portfolio Selection Model. He provides consistent analyses towards optimal combinations of risky assets and risk-free assets, identifying the efficient set of portfolios or the efficient frontier of risky assets that minimizes the variance for any target expected return.

The expected return rate increases proportionally to the level of risk incurred by investors – the price of risk -, yet investors can adjust their preferences combining the optimal risky portfolio with the riskless asset – the price of time -, thus avoiding the diversifiable part of the total risk. The power that this extensive diversification strategy has to eliminate part of the nonsystematic risk is empirically demonstrated by Lintner (1965) and later by Statman (1987). The former quantified the effects of portfolio diversification, using data on NYSE (New York Stock Exchange) stocks. He found that, on average, the risk of equal weighted portfolios of randomly selected stocks decreases significantly with diversification[5].

Lintner (1965)'s seminal research left some relevant recommendations to individual investors and mutual funds managers concerning expected returns and risk. Lintner testified that the best diversification strategy is the one that provides the highest ratio of the expected excess return to standard deviation. Thus, the optimal portfolio resulting from the best combinations of expected return and risk will provide the highest compensation or additional return per unit of extra risk incurred. The obtained results revealed that 57 of the 70 funds analyzed had a higher ratio of expected excess return to risk than the S&P 500, even though the majority of funds average lower returns.

The benefits of diversification were verified through the analysis of the residual risks, measured as the standard deviation of the estimate, obtained by regressing funds returns on the return of the market index (S&P 500 index). The results suggest that, apart from the market risk, the residual risk still accounts for a significant part of the total risk, even in situations of well-diversified and professionally managed portfolios[6].

Individual investors are highly recommended to pursue a prudent and effective diversification strategy to reduce the risk associated with the variability of the expected returns (which are significant, especially in case of common stocks), and to improve the expected return-risk relationship. They should also notice that the lower the correlation between the assets, the greater the gain in efficiency.

---

[5] The estimated average standard deviation of returns of portfolios comprised of only one stock was about 49,2%, and it declines sharply to the minimum of 19.2%, as the number of stocks included in the portfolio increased.

[6] The residual risk was higher than the return of the risk-free asset.

Consistent with the classical CAPM tenets, Sharpe, (1964) demonstrated that the required rate of return of an asset converges to an equilibrium point that is linearly represented as a function of the asset's return rate responsiveness to changes in economic activity (systematic risk or market beta), plus the pure interest rate. Hence assets that are not affected by changes in the economic activity will provide the return of the risk-free asset and those which are sensitive to changes in the economic activity yield on average higher expected return rate. These findings lead to an important premise of the traditional capital market theory by showing that the aggressive securities or those which are more sensitive to market shocks are commonly priced above those whose response is less significant to changes in the economy (defensive securities). As a result, investors demand a higher expected return for investing in aggressive securities than they would require for investing in defensive securities. Which is to say that under the CAPM equilibrium conditions, investors can obtain the highest expected return regardless the level of risk, combining the least risky investment (i.e. U.S. Treasury bills), which has a beta of zero, and the riskier investment, the market portfolio of common stocks (beta is equal to one).

As investors are risk-averse, their willingness to take some risk is compensated with the excess return they require to invest in the market or the market risk premium, which is given by the expected market return minus the risk-free rate. The linear relationship between the expected return on investment and the market risk (beta) is commonly referred to as the Security Market Line, which is a straight line that starts at the point where beta is zero (investment in a risk-free asset). Then, the expected return on risky investments increases proportionally with the market risk – the risk that investors cannot avoid or mitigate through diversification.

In short, CAPM predicts that the rational investors aware of their aversion to risk, will hold the market portfolio or a basket of securities that mimic the market index, in order to obtain the highest level of diversification possible. Therefore, the rate of return required by investors is the contribution of each security to the overall risk of the market portfolio. Analytically the expected risk premium on stocks can be expressed as follows:

- $E(risk\ premium\ on\ stock) = beta * E(risk\ Premium\ on\ market)$

Investors require a premium for holding a risky asset. This premium (or compensation) demanded by investors varies in direct proportion to beta, which measures the extent to which returns on the stock and the market move together.

Black et al. (1972)'s research empirically tests the implications of the CAPM on security pricing. They found disruptive results concerning the relationship between the expected risk premiums on individual assets and their systematic risk or beta. The linear relationship between an assets' return and their betas theoretically established by the CAPM does not always hold. Black et al. (1972) found that the expected excess return of an asset is not strictly proportional to its beta. They built a two-factor model relaxing one of the most important assumptions of the conventional CAPM, which states that investors have unlimited access to riskless borrowing and lending. In equilibrium, the two-factor model predicts that the expected excess return on a security is then given by the expected excess return on a portfolio that is uncorrelated with the return on the market portfolio (zero-beta portfolio), plus the market risk premium subtracted by the zero-beta portfolio, all multiplied by the security's level of systematic risk. The estimated parameters were efficient and statistically significant, thus contradicting the traditional form of the CAPM. Further results indicated that the excess returns on high-beta securities are lower than the expected risk-adjusted return predicted by the CAPM[7]. For a low-beta portfolio of stocks, the estimated excess return was higher than the risk-adjusted return.

Gibbons and Ferson (1985) hypothesize that conventional financial valuations models have failed to analyze the behavior of conditional expected returns over time and incorporate the changes in expectations into their models. The assumptions of constant risk premiums were relaxed, and the refined model suggests that conventional financial valuations models can derive consistent and testable estimators of expected returns without observing the market portfolio or specify the state variables.

The immediate concern of our first model is to analyze the performance of ETFs that track the most popular European and American market indexes and determine whether they can beat the market. To analyze ETFs' ability to the beat the market or to generate

[7] Under the CAPM equilibrium, the expected excess return of an asset is given by the excess return on a market portfolio multiplied by the asset's expected return rate responsiveness to changes in the economy (systematic risk or beta): $E(\tilde{R}_j) = E(\tilde{R}_M) \times \beta_i$.

returns that are higher than the risk-adjusted returns predicted by CAPM, we will take Jensen (1968)'s approach. Jensen (1968) was the first to investigate and systematically test the performance of mutual funds, methodically examining their ability to beat the market. The results show that the funds, in general, were unable to perform sufficiently better than the market or even "beat the market". However, some funds exhibited stunning performance and surprisingly high returns given their level of risk. The best fund in terms of performance yields an alpha of 0.058, exceeding the author's expectations, and regarding their sensitiveness to market changes, the estimated beta for the average fund was approximately 0.85, indicating that most funds were less risky than the market index.

The empirical model underpinning Jensen (1968)'s study is easily derived and it can be constructed in a straightforward way. The parameters of the model can be estimated through OLS (Ordinary Least Squares) and its consistency and statistical significance can be tested using a simple "t-test". Algebraically, it can be defined as:

- $R_{it} - R_{f_t} = \alpha_i + \beta_i(R_{m_t} - R_{f_t}) + u_{i_t}$

This equation is a clear proxy of the CAPM, where the return on security i (R_{it}) discounted from the risk-free rate (R_{f_t}) or the expected excess return is given by the market risk premium adjusted ($R_{m_t} - R_{f_t}$) multiplied by its level of risk (β_i), plus the error term, u_{i_t}.

To determine whether funds can perform significantly better or worse than its benchmark we need to analyze the signal and the significance level of the alpha parameter. Intuitively, the parameter alpha famously known as Jensen's alpha, signals if the fund can earn significant abnormal returns, "in excess" of the market-required return, given the fund contribution to the overall risk. Securities that are able to provide an expected excess return rate higher (lower) than the risk-adjusted return given by CAPM exhibit a positive (negative) alpha.

The following expression clarifies the meaning of risk-adjusted return inherent to CAPM and it helps us understand why increasingly more investors are resourcing to financial instruments that provide "abnormal returns" or positive alpha.

- $R_{it} - R_{ft} = \alpha_i + \beta_i(R_{mt} - R_{ft}) + u_{it} \leftrightarrow \alpha_i = R_{it} - R_{ft} - [\beta_i(R_{mt} - R_{ft})]$;
- $\alpha_i = (R_{it} - R_{ft}) - [\beta_i(R_{mt} - R_{ft})]$.

If the alpha coefficient is positive the expected excess return on fund i $(R_{it} - R_{ft})$ is automatically higher than the risk adjusted return $\beta_i(R_{mt} - R_{ft})$, which means that the reward to investors is even higher than the theoretical compensation required for bearing the risk. And it is a clear indication that the chosen fund outperforms the broad market index, and therefore investors would want to buy it, as suggested by Ferson Lin (2014). Ferson and Lin (2014) contended that a more efficient measure of alpha based on investors preferences can provide reliable buy or sell indication. The results indicate that funds whose client-specific alpha is expected to be larger than the traditional alpha, tend to experience larger inflows. This paper contradicts the existing literature, as most of the studies on this field concluded that traditional alphas are not reliable proxies of the attractiveness of an investment. Therefore, instructions to sell or buy provided by the alpha parameter are not precise nor conclusive.

Chen and J. Knez (1996) found that the client may be willing to short the asset even if it has a positive alpha. Contrarily, Glode (2011)'s paper suggests that clients may be willing to buy a fund even if it has a negative alpha.

2.2.4 From Factor to index Models – An investment Strategy Perspective

We can say that an investor's risk of holding a stock over any given time is given by "how good or bad their expectations were in contrast with reality". Or we can technically say that the risk of holding a stock is that the provided dividends plus the final stock price will be higher or lower than expected. Thus, the capital gains due to favorable price fluctuations and the effective realized returns on an investment is a risk and not a certainty.

Nowadays, it is very clear that market wide news and oscillations in the "state of the economy" can dramatically affect the Stock Market, triggering, for example, downwards/upwards movements in the stock price, as the economy experiences periods of recession/expansion. Central bank announcements about interest rate and political instability can also trigger fluctuations on stocks' prices, and ETFs are no exception as their market quotes can suffer changes throughout trading days, due to unfavorable or favorable news about the economic system or specific firm's news that compose the ETF.

N. Chen, Roll, and Ross (1986) found that traditional sources of systematic risk, such as the spread between long and short interest rates, expected and unexpected inflation, and high-low grade bonds spreads are significantly priced in the market. While potential sources of risk driven by shocks in the oil prices or unexpected changes in aggregate consumption or broad market index are not rewarded in the stock market.

Oscillations of stock returns due to firm-specific news are independent risks, these risks do not represent major problems for the overall Stock Index, as they are unrelated across stocks, not affecting all stocks. They are commonly referred to as firm-specific risk, idiosyncratic risk, or diversifiable risk. On the contrary, market-wide news affects all stocks simultaneously and they cannot be eliminated through diversification. It generates shocks that are correlated across securities, commonly referred to as systematic risk.

The key assumption underpinning the tenets of factor and index models hails from the differences in responsiveness level to macroeconomic shocks among securities. It is reasonable to assume that some securities will be more sensitive than others to economic shocks. In a similar vein, some firms are more capable, both financially and operationally, to surpass and recover from crises at a higher pace than others. Zvie, Kane, and Alex J. (2013)'s book exemplified that auto firms probably will respond more dramatically to changes in general economic conditions than pharmaceutical firms.

Therefore, we can construct our Single-Factor Model, by assigning each firm a sensitivity coefficient that captures systematic risk caused by changes in macroeconomic conditions:

- $R_i = E(R_i) + \boldsymbol{\beta_i m} + e_i$

The rate of return on security i, can be decomposed into the sum of its expected return plus its exposure or sensitivity to unanticipated macro changes (m), and the error term e_i.

Investors can allocate their capital among risky and risk-free asset, investing a fraction of their income in Treasury bills and the remaining part of their wealth in risky assets (i.e. Apple's stock). The expected return resulting from these combinations, plotted as a function of the risk of the portfolio (standard deviation of the portfolio) can be represented graphically as a straight line that starts at the point where the standard deviation of the portfolio is zero (intercept – return of risk-free asset). And then increases proportionally with the level of risk. The slope of the expected return of the portfolio represents the increase in the expected return of the complete portfolio per unit of additional risk/standard deviation – commonly mentioned in the literature as the incremental return per incremental risk or reward-to-volatility ratio. The straight line connecting all the feasible pairs of risk-return combinations represents the Capital Allocation Line.

Defenders of a passive strategy argue that passively held risky assets are highly valuable for investors seeking cost-effectiveness, avoiding expensive direct or indirect security analysis, high exposure to the volatility of actively managed stocks, as well as the need for meticulous and highly frequently forecasting and speculative analysis about the stock market.

Malkiel (2003) provides empirical evidence in favor of the passive investment strategies for all types of investment markets. Pástor, Stambaugh, and Taylor (2015) show that active mutual funds return decrease with the size of the industry. Their abilities to outperform the benchmark depend negatively on the size of the active mutual fund industry.

Active investing, as the name suggests, requires some "work", and a "hands-on" approach, daily. Actively managed portfolios are subscribed or assigned to a professional money manager who is responsible for performing a wide range of security "scanning" as well as quantitative and qualitative analyses, with the objective of anticipating and reacting properly to short-term price fluctuations.

Active investing supporters contend that meticulous and constant security analysis, as well as well-developed speculative strategies, can lead investors through highly profitable paths, exploiting market imperfections and mispricing opportunities, during periods preceding the automatic price readjustment and convergence to equilibrium.

Unlike passive investors, active investors do not face any source of limitations about trade frequency or the amount of buying and selling within their portfolio, as they can and should buy or sell whenever they think it is the right time to do so.

Another argument often summoned for this discussion is tax-efficiency. Usually, passive investors can benefit from low tax charges on capital gains due to their long-term buy-and-hold strategy.

However, passive strategies receive much criticism due to their limited and restricted investment plan and due to their small return investment scale, as the majority of passive funds are limited to a specific index or predefined set of investments with little volatility as possible. Therefore, a passive investor is locked to those holdings, regardless of the changes that might occur in the market, thus exhibiting in most cases, steady and very rigid growth patterns.

Active managers on the other hand, have no limitations to buy or sell and reallocate their capital in harmony with current market "sentiment", using hedging strategies to reduce the risk that they face from potential future movements in a market variable, and other financial techniques like short-selling and put options to minimize their risk, and increase the rewards (see, Cheng, Jegadeesh, and Wermers (2000)).

An investor can choose to delegate the managerial task or the responsibility to perform a security analysis to the care of a professional or decide to invest the time and cost to acquire all the necessary information to generate or build from scratch an optimal active portfolio of risky assets. In either case, he must be aware of the commissions and managerial fees that will be charged. For this reason, the constitution of an active portfolio is more expensive than a passive one.

Brinson, Hood, and Beebower (1995) illustrate that, on average, active management fees cost the average investment plan 1.10% per year. In turn, passive investors devote no resources or time to acquiring information on any individual stock or asset, which will form the portfolio. A plausible and highly recommended passive strategy is the "Neutral Diversification Strategy". Under this strategy, investors must form a well-diversified portfolio of common stocks. One possible way to construct a diversified portfolio of stocks is to select a group of stocks that mirror the value of the corporate sector of the economy, for example, passive investors can choose a portfolio of stocks that replicates or mimics the overall performance of the stock index (i.e. S&P 500).

The use of the index model to construct risky portfolios was originally developed in L. Treynor and Black (1973). The Single-Index Model can be derived through a simple modification or operational improvement of the single-factor model, if the rate of return on a broad index of securities, such as the S&P 500, can be used as a valid proxy for the common macroeconomic factor.

This approach is very useful and practical, as we can easily extract a large amount of data and information about the market index, and estimate its volatility, performance patterns and its correlation as well as its similarities in terms of behavior and returns, with individual securities.

The regression equation can be expressed as follows:

- $R_{i_t} = \alpha_i + \beta_i(R_{M_t}) + e_{i_t}$.

The excess return of security i at time t is denoted by R_{i_t}. The Greek letter alpha (α_i) represents the security's expected excess return when the market excess return is zero, that is the observable excess return of a stock when the market is neutral, it is commonly referred as the nonmarket premium. On the other hand, the slope coefficient, β_i, measures the security's sensitivity to the market index. It basically tells us by how much the security return tends to increase or decrease for every 1% change in the return of the index.

The error term or the residual (e_{i_t}) accounts for the firm-specific surprise or sudden events or news that are intrinsically related with current or future activity of the individual firm.

Another important item worth mentioning, regarding risk and covariance in the Single-Index Model, is the role played by security betas in determining dispersion and joint variability measures – variances and covariances respectively.

The total risk in the single-index model can be decomposed into systematic risk and firm-specific risk, the former depends exclusively on security's sensitivity to shocks in the market index area (β_i) and the overall market risk (σ_M^2).

Uncertainty about general economic conditions, such as interest rates, inflation or Gross National Product (GNP), are clear examples of systematic risk, as it affects many assets, asymmetrically or in a not homogeneous or uniform way since each firm reacts or responds differently to market shocks. Nonetheless, firm-specific risk ($\sigma_{e_i}^2$), depends only on firm-specific news or events, triggered by announcements that do only concern the firm's core activities, business model or financial decision (e.g. announcements about new technology and highly sophisticated hardware and chips developed by Apple).

Investors should keep in mind that "same economic shocks or changes in the macroeconomic conditions can affect, for example, Apple and Supernus Pharmaceuticals[8] stocks' prices differently". Thus, each firm exhibits a specific Beta, which captures changes in its expected rate of return due to unanticipated changes in general macroeconomic conditions.

I end this section reinforcing the idea, that the "magnitude or intensity" of security's reaction or responsiveness to macroeconomic shocks captured by broad market index, tends to be different across securities. Some stocks are more prone to incorporate overall market changes, as compared to others.

2.2.5 Alternative Models – A Modern Approach

The seminal research brought about to the academic atmosphere by Fama and French (1993), and Carhart (1997) changed the perception of former factor models and other market models devoted to one single manifestation of systematic risk.

The Three-Factor model introduced by Fama and French, which was later complemented by Carhart's Momentum Factor, clearly undermines the effectiveness and practicality of CAPM, as it recognizes that the market index alone cannot fully explain the variation in asset returns. There might be factors other than the market portfolio that entails changes in security's return. For better estimation and improvement in traditional asset pricing mechanisms, we need to recognize that securities are exposed to multiple sources of systematic risk. Allowing for several risk factors we can achieve better descriptions and estimations of security returns. Unlike the CAPM, the Arbitrage Price Theory (APT) is open to multiple sources of risks. Ross (1976) contend that a multifactor model can provide estimations that are more closely related to reality.

Despite the differences, APT and CAPM are very similar in terms of the risk-return approach. Both end up highlighting the distinction between the systematic or non-diversifiable risk, and the diversifiable risk, which contrarily from the diversifiable risk does not require a reward in the form of a risk premium.

[8] Supernus Pharmaceuticals (SUPN) is a quoted pharma company, specialized in drugs for treating diseases of the nervous system.

In the meantime, the underlying theory and the underpinnings of APT depend on the assumption that a rational equilibrium in capital markets rules out any kind of arbitrage opportunities, resulting from the systematic exploitation of security mispricing or other sources of imbalance in security prices. Furthermore, the expected return-beta relationship brought about by APT theory requires a well-diversified portfolio that can be constructed from scratch using a large number of securities. CAPM takes instead a quite different and predominant risk-return approach, stating that every investor aims at holding the optimal risky portfolio, under the light of the mean-variance efficiency theory developed in Markowitz (1952). This strategy implies that the expected return-beta relationship should hold for all securities, contrary to the APT which argues that this relationship holds for a restricted number of securities.

Fama and French (1993) investigates the role played by specific firm-characteristics to explain stocks' average returns. They demonstrated that size and book-to-market equity are highly related to asset returns and the prominent source of risk. Nowadays, the Three-Factor model plays a major role in the capital markets ecosystem and its pertinence to the real-world industry is obvious. It starts with a simple assumption that the market index *per se* cannot fully explain all the variation in asset returns, questioning that there might be other sources of risk affecting securities' returns.

The theoretical underpinnings of Fama and French's findings hails from the classical tenet of financial theory. Which relates stock's fundamental attributes to their expected returns. Typically, some types of stocks yield, on average, higher returns than others. Small companies value stocks or those with low P/E ratio, and commonly undervalued by investors who do not recognize their full potential and the worth of their fundamentals (e.g. sales, earnings), typically yield higher returns than growth stocks (high P/E and promising future ahead, according to investors beliefs).

Stocks that are experiencing a bright moment described by systematic increases and appreciation in their share price tend to generate higher returns than the ones exhibiting a steady and significant period of negative returns and depreciation of its market value. Therefore, it is undeniable that firms or portfolios' characteristics are very likely to be related to security's returns. However, the main question here, is how we can derive or trigger the risk factor component out of these variables, size and book-to-market ratio (BE/ME). Fama and French (1993) provide consistent evidence and signs of the persistence of assets' earnings due to firms' specific characteristics, such as book-to-market equity.

Firms with high BE/ME (low stock price relative to book value), exhibit a tendency of low earning on assets, which persist for at least five years before and five years after BE/ME is measured. Regarding the size factor, Fama and French (1993) conclude that small firms tend to have lower earnings on assets than bigger firms. The size effect on firms' earnings turns out to be more prolonged and persistent during periods following the 1980s recession. Especially for small firms, earnings on assets were completely devastated by the recession, comparatively with big firms. Thus, small firms are exposed to longer earnings depression than large firms.

We can say that the BE/ME and the Size factor are related with common risk factors that might be useful in explaining the average stock returns. The results achieved through this study can be widely applied to any research that requires meticulous estimation of expected stock returns, such as seminal studies related to performance analysis, portfolio selection, or even papers dedicated on measuring the abnormal return.

The common risk factors introduced by Fama and French (1993) did a great job explaining variation in stock returns. They contributed significantly to the extant literature on asset pricing providing three major potential candidates for relevant sources of systematic risk.

The following section describes the data and presents the methodology and the econometric setup employed.

3. Data, Methodology, and Econometric setup

3.1 Data

Our sample is composed of 18 ETFs that track major European and American Indexes. We use daily historical data from March 3, 2005, to December 30, 2017. For each ETF, we gather data about price, volume, closing bid and ask prices, price-to-book ratio, net asset value, market capitalization and the historical closing price of the underlying benchmark. For this study, we use daily data as in past studies (see Peltomäki (2017)) as it is more adequate because the creation and redemption features of ETFs, which normally take place at the end of the day to meet the daily pricing pressure, thus making daily data particularly interesting for ETFs.

The data were obtained from the Bloomberg terminal and we set the euro as the currency measure. After collecting the data, the final sample was organized according to the following criteria. First, we only select ETFs that track or replicate explicitly a benchmark stock index. Therefore, we focus exclusively on index-tracking ETFs.

Second, for each index, we select the ETF with the most extensive data available. As ETFs are a relatively recent phenomenon some ETFs were excluded throughout this process due to low data availability.

3.1.1 Summary Statistics

Table 1 reports the summary statistics of the entire sample and the description of the main variables used in this study.

<Table 3.1 Summary Statistics of the sample and Description of Variables>

The results show that the average daily percentage return of the sample's ETFs is positive and equal to 2.3 basis points(bps) or 0.023%, and the corresponding benchmark indexes exhibits an average daily return of 2.5bps (0.025%). The overall daily average returns on 18 ETFs underperform those of their benchmark indexes. The average market risk premium is also higher than the average risk premium provided by the sample's ETF, at daily basis. The sample's spread and spread ratio averages 0.331€ and 0.6%, respectively.

The number of shares traded on average per day amounts to almost 16 million shares for the entire sample, and the average turnover is close to 44.6 thousand of shares. The average NAV of the 18 ETFs sample is above 100€ and the average market capitalization exceeds 34 Billion of euros.

As reported in **Table 2** the number of shares traded, as well as the turnover is significantly higher for ETFs tracking U.S. indexes comparatively with European ETFs. As we might expect differences in the market value and the net worth of the assets backing up the fund are very significant as well, due to the dimension, magnitude, and deepness of the U.S. market. ETFs tracking the three largest stock indexes in U.S. (S&P

500, Nasdaq and DJIA) exhibit the highest volume, turnover and market value, while the BELL 20, TDT NA, and the PPP ETFs exhibit the lowest volume, turnover and market value, on the average daily basis. The fundamental value of an ETF or the value of its underlying securities reflected in the NAV price is also higher for ETFs tracking the major U.S. indexes.

<Table 3.2 Fund's Characteristics>

The ETF tracking the performance of the most important stock indexes in Spain (IBEX indexes) offers an average spread of approximately 95 cents per share, which is somehow 90 cents higher than the average spread charged by the QQQ ETF (the lowest average spread ETF). The spread changes significantly among ETFs, meaning that the trading activity is different across ETFs. A relevant information to investors that we can take from here is that some ETFs are more difficult to trade than others, even for large-scale and well diversified ETFs, the additional cost due to bid-ask spread can constitute a real problem to investors that trade actively a large number of shares.

Meanwhile, the returns due to changes in the net asset value of ETFs or the NAV return are very similar from the returns due to changes in ETFs prices, therefore we found no significant evidence of deviations of ETFs prices from its NAV.

ETFs tracking the Portuguese and the Italian market indexes have the lowest average daily return of -1.9 (-0.019%) and -0.21 (-0.0021%) basis points, respectively. ETFs tracking the CRSP U.S. Large Cap Value Index and CRSP U.S. Total Market Index exhibit the highest average daily return of 4.9 (0,049%) and 4.5 (0.045%) basis points.

The average risk premium for most ETFs tracking European indexes is close to 0.24%, and for the U.S. market, the required risk premium demanded by investors is a bit higher, somewhere close to 0.28%. The largest ETFs linked to the U.S. market, QQQ, VTI and the SPY ETFs are expected to earn on average a return of 5bps (0.05%), 7bps (0.07%), and 3.5bps (0.035%) above the return rate of the risk-free asset. On the other hand, investors demand on average an extra return or a premium rate of 3bps (0.03%) and 2bps (0.02%) to bear the risk of the portfolios of stocks matching the performance of the iShare Russell 2000 and CRSP U.S. indexes (the smallest ETF included in our American subsample of ETFs).

For European ETFs the results are slightly different, ETFs tracking the German, Swiss and Belgium market stand out with an average expected the excess return of approximately 6bps (0.06%), 5bps (0.05%), and 4,5bps (0.045%), respectively. The ISF LN ETF that tracks the performance of an index composed of the 100 largest UK companies, provides a relatively low-risk premium of 2bps (0.02%). The average excess return of the ETFs tracking the Portuguese, Spanish and the Italian market are very close to 2bps (0.02%), a relatively modest performance compared with their peers.

In more detail, the results presented in **Table 3** show that most of ETFs tracking American stock markets exhibit higher average returns and lower standard deviations than ETFs tracking European benchmark indexes.

<Table 3.3 Model 1 Descriptive Statistics>

This evidence of superior return is confirmed in **Table 4** where the results are tabulated by the market. Panel A shows that the average excess return of our European ETFs' sample is around 3.7bps (0.037%) while the excess return of our U.S. sample of ETFs averages 4.3bps (0.043%), thus ETFs linked to the performance and volatility of the major European market indexes underperforms American ETFs by an amount of 0,6bps (0.006%).

Consistent with previous findings **Table 4** revealed a certain degree of underperformance of ETFs compared to their inherent benchmark indexes. Panel A and B report that the European and American subsamples of ETFs underperform their benchmark by 0.3bps (0.003%) and 0.2bps (0.002%) respectively.

The results for the liquidity measures suggest that the average daily spread, spread ratio, and Amihud illiquidity factor are higher for the European market. The highest daily spread for the European and American subsamples of ETFs amounts to 6.22€ and 5,55€ respectively, conversely, the minimum spread equals to zero for both markets, which is justified by the trading periods where the highest price offered by the buyer equals the minimum acceptable price for the seller. The natural logarithm of volume and turnover are higher for the U.S. market, comparatively with the European market on the average daily basis.

<Table 3.4 Model 2 Descriptive Statistics>

The excess return on ETFs and the risk premium provided by the market exhibit a very high correlation, around 98%, as demonstrated in **Table 5**. Evidence of a high-level correlation between ETFs and market indexes suggest that the major goal of ETFs as tracking instruments is being accomplished.

<Table 3.5 Correlation Matrix>

The excess return on ETFs exhibits a significant and positive correlation with the size and illiquidity factor, the value factor is negatively correlated with ETFs excess return. Spread, volume and turnover exhibit a small and negative correlation with the excess return rate of ETFs.

3.2 Methodology

3.2.1 Measuring the Abnormal return of major European and U.S. ETFs

To answer properly the question concerning the ability of major European and American ETFs to generate an abnormal return, we adopt the Jensen's Alpha approach and inherent methodology, so greatly referenced in the financial academic-circle. The methodology and technical specifications behind Jensen (1968)'s seminal study are quite simple and straightforward.

Jensen (1968) uses a simple hypothesis test method, commonly known as t-test to evaluate the ability of mutual funds to outperform the market. To test whether ETFs can beat the market, I'm going to employ the Single-Index Model discussed in the previous chapter, and therefore perform simple statistical inference to check on the significance of the intercept parameter, alpha.

The underlying hypothesis is that the European and U.S. subsamples of ETFs can generate returns that exceed the risk-adjusted return predicted by an equilibrium model such as CAPM, thus the alpha parameter should be positive. This means that the expected rate of return generated by our European and American subsamples of ETFs is even higher than the compensation initially required by investors for investing in that investment vehicle, given its level of risk.

Our first model will take the form of:

$$\textbf{(1)} \quad r_{i_t} - r_{f_t} = \alpha_i + \beta_i\left(r_{m_t} - r_{f_t}\right) + u_{i_t}$$

Where: $r_{i_t} - r_{f_t}$ – is the return of security i (ETFs that tracks majors European and American market index), at time t, minus the return of the risk-free asset (U.S. Treasury Bill 10-year yield change rate) – the expected excess return on ETFs ; $\beta_i\left(r_{m_t} - r_{f_t}\right)$ – is the fraction of return due to movements in the overall market, it reveals the investors' "appetite" for the market risk or the market risk premium $(r_{m_t} - r_{f_t})$, given security's responsiveness to changes in the overall market index (β_i).

The coefficient alpha (α_i), is the intercept of our equation, it measures the nonmarket premium or the abnormal return; and u_{i_t} – is the error term or the residual of our model, which is given by the difference between the actual ETFs excess return and its fitted value or the estimated values of ETFs excess return by our model.

The critical point here is not just to run the regression and obtain the parameters, but to verify if our model is capable of consistently explaining all the variation in the dependent variable (ETFs excess return) or not. The first thing we can do to ensure that the estimated parameters or coefficients are reliable and significant – meaning that it represents the true relationship between the "regressed" and the predictor variable – is to perform statistical inference, aimed at testing to which extent we can reject our hypothesis of statistical significance of the parameter alpha.

To state matters plainly, the framework or the main scope of our hypothesis testing about the individual parameter alpha obtained from the sample of our bivariate linear regression model can be intuitively expressed as:

- $H_0: \alpha_i = 0$ – Null Hypothesis;
- $H_1: \alpha_i \neq 0$ – Alternative Hypothesis.

The null hypothesis being tested is that the true value of parameter alpha should be zero, against the alternative hypothesis – the true value of alpha is different than zero. The main interest here is to reject the null hypothesis in favor of the alternative hypothesis. Therefore, a positive and significant alpha would suggest, that European and U.S. located ETFs being examined are able to outperform the underlying market index, generating an expected excess return rate that is above the risk-adjusted return or the initial risk premium or compensation required by ETFs' investors.

The statistic to test the hypothesis stated above, is called the t-statistic or t-ratio of the parameter alpha, and it can be easily calculated as:

- $t_{\widehat{\alpha}_i} = \frac{(\widehat{\alpha}_i - \alpha_i)}{se(\widehat{\alpha}_i)} \sim t_{n-k-1}$.

3.2.2 Performance Analysis

To analyze the performance of ETFs we will use some well-known performance indicators, such as the reward-to-volatility ratio, and tracking error. The reward-to-volatility takes many forms. Sharpe (1966) devised what we now call the Sharpe ratio, which is a consistent way of examining the performance or the desirability of an investment strategy by adjusting for its risk, or it can be defined as the extra return that investors demand, to bear a portfolio of risky asset – incremental return per incremental risk.

The analytical expression of the Sharpe ratio is given by:

- $Sharpe\ ratio = \frac{E(r_i) - R_f}{\sigma_i}$.

Where the excess return on ETFs ($E(r_i) - R_f$) is adjusted to the level of risk that the portfolio represents to investor (σ_i).

Another, common measure of the reward-to-volatility ratio is the Sortino ratio developed in Sortino and Price (1994)'s seminal research. It is a clear modification of the Sharpe ratio, but it uses a different measure of risk, as it takes as the denominator the downside deviation, which accounts for the "bad side of risk" or the standard deviation of the negative excess returns of the portfolio.

This improved version of risk-adjusted returns can mitigate the effect of high returns outliers, which is, in fact, a clear limitation of the Shape ratio since it does not distinguish between upside and downside volatility.

Treynor ratio is another version of the Sharpe ratio which strongly disagrees with the risk measure undertaken by Sharpe. Unlike Sharpe and Sortino ratio, Treynor ratio places major weight on the systematic risk of the portfolio, instead of using both (systematic and idiosyncratic risks).

- $\text{Sortino ratio} = \frac{E(r_i) - R_f}{DSD_i}$;

- Treynor ratio $= \frac{E(r_i)-R_f}{\beta_i}$.

The denominators DSD_p and β_p accounts for the downside deviation of the ETF and the market Beta, respectively.

Regarding the tracking efficiency of ETFs, the most common measure present in the literature is the tracking error. Tracking error basically indicates how closely a portfolio follows the underlying benchmark, so it can be either calculated as the absolute difference in returns of the index portfolio and the benchmark index or as the standard deviation of the difference between the portfolio and index returns (see, Roll (1992), Pope and Yadav (1994), and Larsen and Resnick (1998)).

Therefore, the two measures of tracking error can be expressed as follows:

- $TE_{1,i} = \frac{\sum_{i=1}^{n}(r_{it}-r_{bt})}{n}$
- $TE_{2,i} = \sqrt{\frac{\sum_{i=1}^{n}(r_{it}-r_{bt})^2}{n-1}}$

Where:

r_{it} – the return on ETF i in period t;

r_{bt} – the return on the benchmark index, and n is the number of observations.

The next section is devoted to the construction of our second model aimed to test the effect of liquidity on ETFs returns.

3.2.3 Determinants of ETFs returns

Our second model will capture the extent to which European and U.S. index-tracking ETFs deviate from its expected equilibrium value due to unexpected changes in liquidity. To construct our model, we start defining the measure of liquidity that best match investors' expectations about how quickly they can buy or sell ETFs shares without suffering much with price adjustments that affect negatively their ability to trade profitably.

Regularly most of the non-institutional investors buy and sell shares of ETF directly from the secondary market, trading ETF shares that already exist. For this specific trading channel what matters the most for investors in terms of liquidity is the information about the volume of shares traded and spreads provided by the broker.

On the other hand, institutional and large-scale investors can perform large trades in the primary market, avoiding price volatility and major disturbances of the secondary market. The supply provided by the authorized participant can also be adjusted to meet in an efficient way their demand due to in-kind creations and redemptions process. That is why the determinants of the liquidity of an ETF traded in the secondary market and the ones traded in the primary can be different. Secondary market liquidity depends exclusively on the value of shares traded, while primary market liquidity is more a function of the value of the underlying shares that back the ETF.

In our study, we decided to use liquidity measures that best reflect the level of marketability of the secondary market. The main source of ETF liquidity that we are going to employ into our second model is the "visible" type of liquidity resulting from trading activities of buyers and sellers in the secondary market. Consistent with the literature (see, Amihud and Mendelson (1986), Ibbotson et al. (2013), and Amihud (2002)) we will use the following liquidity measures:

1 – Average Daily Volume (ADV) – is the average number of shares traded at daily basis. A very important metric of liquidity that basically informs investors about how easy or difficult it can be for them to trade for example 100 shares or 100.000 shares of Comstage PSI20 UCITS (PPP) in a single trading day;

2 – Bid-Ask spreads – is the difference between the price at which someone is willing to buy a certain security and the price at which someone is willing to sell it.

3 – Spread ratio – is the closing bid-asks spreads expressed as a percentage of ETFs prices;

4 – Amihud Illiquidity factor – is given by the ratio between daily return and daily dollar trading volume. This ratio represents the percentage price change per dollar of trading volume. We denote the Amihud, (2002) illiquidity factor as $ILLIQ_{iy}$ and it can be expressed as follows:

- $ILLIQ_{iy} = \sum_{i=1}^{D_{iy}} \frac{|r_{it}|}{Dvol_{it}}$

Where, r_{it} and $Dvol_{it}$ stand for the daily return and daily dollar trading volume for security i at time t, respectively.

5 – Turnover – is defined as daily trading volume divided by the end of day shares outstanding.

To complete our model, we use the mimic size and value factors developed in Fama and French (1993)' seminal research. The mimic size factor was computed as the difference between the daily average returns of the three smallest ETFs (three small-ETFs portfolios) and the daily average returns of the three biggest ETFs (three big-ETFs portfolios).

The mimic return factor related to Value is given by the difference between the daily average return of the three ETFs portfolios with the highest book-to-market equity and the daily average return of the lowest book-to-market equity portfolio of three ETFs.

Algebraically it can be illustrated as follows:

- Size (Small minus Big) = $\frac{\sum_{i=i}^{n} r_{iS}}{3} - \frac{\sum_{i=i}^{n} r_{iB}}{3} = \frac{(r_{1S}+r_{2S}+r_{3S})}{3} - \frac{(r_{1B}+r_{2B}+r_{3B})}{3};\ (n = 3);$

(r_{S1}) is the return of the first ETF belonging to the portfolio of the three smallest equity ETFs, r_{2S} and r_{3S} is the return of the second and third ETF that composes the portfolio of the three smallest ETFs, consecutively. In turn r_{1B}, r_{2B}, and r_{2B} is the return of the first, second and third ETF, composing the portfolio of the three biggest equity ETFs, respectively.

- Book-to-Market equity (High minus Low) = $\frac{\sum_{i=i}^{n} r_{iH}}{3} - \frac{\sum_{i=i}^{n} r_{iL}}{3} = \frac{(r_{1H}+r_{2H}+r_{3H})}{3} - \frac{(r_{1L}+r_{2L}+r_{3L})}{3}, (n = 3)$

r_{1H}, r_{2H}, and r_{3H} stands for the daily return of the first, second, and the third ETF on the three high-book-to-market (low-price-to-book ratio[9]) equity portfolio, respectively. On the other hand, we have the daily return of ETFs belonging to the three lowest book-to-market equity (highest price-to-book ratio) portfolio – r_{1L}, r_{2L}, and r_{3L}.

Then, the second model of this research created as an attempt to capture the impact of liquidity on the expected excess return rate of ETFs that tracks major European and U.S. market indexes will take the form of:

$$\textbf{(2)}\ R_{i,t} = \alpha_i + \beta_M(R_{M,t}) + \beta_{Size}SMB_t + \beta_{value}HML_t + \beta_{Vol}Log\ vol_t + \beta_{turn}Log\ turn_t + \beta_{s1}Spread_t + \beta_{s2}Spread\ ratio_t \pm \beta_L ILLIQ_t + \varepsilon_t$$

[9] For each fund, the P/B ratio is computed as the weighted average of the P/B ratios of all stocks composing the fund.

The excess return of a security i at time t, is estimated based on: the intensity of security's responsiveness or reaction to market changes captured by $\boldsymbol{\beta_M}$, the mimic return factor related to size and value ($\boldsymbol{\beta_{Size}}$ and $\boldsymbol{\beta_{value}}$), and the liquidity measures described above, plus the error term ε_t.

3.3 Econometric Setup

The data sample is arranged in panel data form because our dataset comprises both time series and cross-sectional elements. The ETFs constitutes the entities upon each we will perform separate and market-based regression analysis. ETFs are then sorted by market. The European subsample is composed of all ETFs that track European benchmark indexes and the U.S. subsample aggregate all the ETFs that replicate major benchmark indexes in the U.S.

We perform several tests to address issues concerning model specification and frequent problems with long time series data, such as heteroskedasticity, instability in distributions, and serial correlation.

The underlying null hypothesis of the White's test for heteroskedasticity was strongly rejected, therefore we have clear evidence that the variance of the residuals varies systematically with the known predictors, thus generating heteroscedastic estimators (as shown in **Figure 3.1**).

<*Figure 3.1 Plot of Squared Residuals and Linear Prediction*>

In the presence of heteroscedasticity, OLS estimators will produce unbiased coefficient estimates and misleading inferences. To avoid that, we employ robust standard errors estimates, which are modified by conception to account for the heteroskedasticity.

The results of the Wooldridge test for autocorrelation in panel data suggest that our sample has a first-order correlation in residuals. The Fisher-type unit root test based on augmented Dickey-Fuller test reveals that at least one panel is stationary, thus the null hypothesis that all panels contain unit roots is strongly rejected.

Given the fact that we need to control for the effect of time-invariant characteristics across our sample of ETFs, we must decide between fixed or random effects models. We run the Hausman test where the null hypothesis is that the preferred or the more adequate model given our sample is random effects against the alternative hypothesis, fixed effects. This test basically checks whether the unique errors are correlated with the regressors or not.

In favor of the null, the results suggest that random effects are the more appropriate model, meaning that it is capable to produce more efficient estimation than the fixed effects approach.

We also estimate feasible estimators through GLS (Generalized Least Square) model to account for the presence of heteroscedasticity and one order autocorrelation in residuals.

The relationship between liquidity and assets returns could be nonlinear. Amihud and Mendelson, (1986) documented a concave relationship between asset return and spread, by showing that the liquidity premium should increase with the bid-ask spread at a decreasing rate.

Therefore, in our robustness checks analysis, we use square variables of spread, spread ratio, and Amihud illiquidity measure to test the impact of nonlinear parameters on the ETFs excess returns.

The interaction term (illiquidity times spread) allows us to assess the combined effect of illiquidity or spread on ETFs excess return. By adding the interaction term, the unique effect of illiquidity on the dependent variable would not be limited to a single coefficient, but it would also consider the value of the coefficient of spread.

In these terms, the effect of illiquidity on ETFs excess returns will also depend on the variability of the spread. We also included a lagged value of the dependent variable, so we can address the impact of one-period lagged ETFs excess returns on the current excess returns.

We introduce 2 ETFs dummy variables to investigate the role played by the two most prominent ETFs in terms of market liquidity, dimension, and investor’s preferences. The first dummy ETF is the SPY ETF which tracks one of the deepest and largest market indices in the World, the S&P 500 index. The second is the C40 FP ETF which tracks the CAC40 index.

The SPY ETF was normally selected due to the dimension and attractiveness of the S&P 500 index, while the C40 FP ETF was chosen because it tracks one of the most liquid market indexes in Europe concerning innovative and leading-edge financial products.

The returns are calculated as the difference between the most recent closing price and the last closing price all divided by the last closing price. For estimation purposes, we use the continuous return also denoted the log returns[10], due to the benefits of normalization and consistency over long time series. In a similar view, to model the price evolution of ETFs, Volume and Turnover are also transformed into logarithmic variables.

To improve the quality of data and avoid noisy observations, we discard the first month of data and extreme values are replaced at 1% and 99% percentiles.

4. Results

4.1 Testing Abnormal return

Panel 1 of **Table 8** shows that the estimated beta coefficient for both markets is statistically significant (at 1% level) and close to one, which confirms the sensitiveness of ETFs to shocks in the market. Index-tracking ETFs are designed to move alongside with their inherent benchmark indexes, therefore the beta coefficient is expected to be close to one.

<Table 4.1 Model 1 Testing Abnormal Return Tabulated by Market>

The estimated Jensen's alpha parameter is negative and very close to zero for both markets, suggesting that the ETFs included in our sample tracking major European and American benchmark indexes were unable to beat the performance of the underlying indexes.

Additionally, we perform a mean comparison test to evaluate the extent to which the return on ETFs differs from the return provided by the benchmark indexes. We found no significant evidence of superior or inferior performance, the existent differences in returns are not statistically significant (t-stat are very low), see Panel 2 of **Table 8**.

[10] The continuous returns are calculated as the natural logarithm of the most recent closing price divided by the last closing price, analytically it can be expressed as: $\ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$.

The results for the individual sample of 18 ETFs, presented in **Table 9**, shows that almost all ETFs exhibit a negative and very close to zero alpha parameter, and a beta coefficient of approximately one.

<Table 4.2 Model 1 Testing Abnormal Return Tabulated by ETF>

Regarding the goodness of the fit, the results are very satisfactory, the proposed explanatory variable (Market risk premium) does explain almost all the variations in the dependent variable. The R^2 of all the regressions are well above 90%, as well as the adjusted R^2 (see **Figure 4.1** and **Figure 4.2**).

The high quality of the linear fit is also present in Milonas and Rompotis (2006)'s research, and it highlights the fact that ETFs are doing a good job replicating the performance of the underlying market, meeting their tracking goals successfully. The good dimension of our sample, an average of 2308 observations per ETF, may also explain the high level of the goodness of the fit.

Following these results, we will redirect our focus to the second goal of this paper, which is to determine the main determinants of ETFs returns and to analyze to which extent oscillations in liquidity can explain unexpected changes in ETFs returns in the two markets.

4.2 Performance Analysis

Table 6 and **Table 7** show that the Sharpe ratio, Treynor ratio, and Sortino ratio are higher for the U.S. sample of ETFs, comparatively with the European sample. ETFs tracking the CRSUP U.S. and the Germany market exhibit the highest Sharpe ratio around 3%, while the LYXIB and the ISF LN ETFs offer the lowest Sharpe ratio, close to 1%.

The Sharpe ratio of all ETFs is positive and greater than one, meaning that the examined ETFs were able to provide to their investors a return greater than the inherent risk they took (the additional increment in the average return earned in excess of the risk-free rate per unit of volatility or risk taken is positive).

<Table 4.3 Reward-To-Volatility Indicators>

<Table 4.4 Reward-To-Volatility Indicators Tabulated by Market>

The Sortino ratio shows some improvements, providing results closer to the overall market expectations about the fair reward-to-volatility ratio of ETFs tied to the U.S. market, which was recently established to be something close to 2%.

The Sortino ratio obtained for our sample of ETFs tracking U.S. benchmark indexes is precisely 3%, contrarily for the European ETFs, the results obtained are slightly below (2.8%).

The Treynor ratio for most of ETFs are quite modest and small regardless the region of the index being tracked, in the meantime, this result is not surprising because Treynor ratio takes only into account the systematic component of the risk (beta), deteriorating the ratio since ETFs are managed to move in the same direction as the market, thus, by design they should respond properly to any changes in the market (Beta close to one), therefore this intentional exposure to the market contributes largely to the overall risk of each ETF.

The major goal of ETFs, which is to track consistently its inherent benchmark, is being partially accomplished, for both European and U.S. subsamples of ETFs. The tracking error is relatively small but not statistically equal zero, therefore, the ETFs included in our sample are doing a relatively good job tracking the performance of the market index, delivering a return close to the return provided by the market, but the tracking error cannot be ignored.

4.3 Determinants of ETFs Returns

Table 10 presents the baseline results of the estimation of the second model. Panel 1 reports our first estimates based on robust OLS regression, Panel 2 reports the random effect estimates, and Panel 3 reports the GLS estimates[11].

<Table 4.5 Determinants of ETFs Returns>

[11] The results of the fixed effect regression presented in **Table 12**, shows that the coefficients of all liquidity variables are not statistically different from zero under a significance level below 10%, which confirms our test that the random effects model is more appropriate.

The results of the robust regression indicate a strong and statistically significant relationship between the expected excess return on ETFs and the excess return on the market portfolio, at a 1% significance level.

As predicted in the CAPM, the expected return-beta relationship holds in our model. For the European sample of ETFs, it ranges from the 0.953 to 0.955, and always 0.999 for the American sample of ETFs (Panels 1 – 3).

The coefficient of mimic value factor is positive, and it ranges from 0.147 to 0.152 for the European sample of ETFs, with a significance level of 1%, and for the American subsample, it ranges from -0.0102 to 0.0240 with little or no degree of significance under the robust and RE specification, but the GLS estimate is statistically significant at 1% level (Panels 1 – 3).

The size factor exhibits a negative relationship with our dependent variable (excess return on ETFs), it ranges from -0.0964 to -0.00989 in the case of European sample, and for the American sample of ETFs it ranges from -0.0608 to -0.04. The robust, GLS and random effect estimations are statistically different from zero, at 1% level (Panels 1 – 3).

The random and GLS estimates of the coefficient spread for the European subsample are negative, while in the case of the American subsample it is positive.

It is statistically significant at least at 5% and 1% level in the case of the European subsample of ETFs, ranging from -0.000207 to -0.000188 (Panels 2 – 3). For the American market, it ranges from 0.000000162 to 0.0000370, but no evidence of statistical significance was found (Panels 1 – 3).

The results for the coefficient of spread ratio are similar since the evidence of statistical significance is only observed in the case of the European sample of ETFs, where the random and the GLS estimates exhibit a 10% and 5% degree of statistical significance.

The coefficient of Amihud liquidity measure exhibit a positive and strong relationship with the excess returns provided by the ETFs that tracks the performance and volatility of European benchmark indexes, it ranges from 0.406 to 1.562, and it is always statistically significant at 1% level. For the American sample of ETFs, it ranges from -0.00841 to 0.0433, but it is not statistically significant.

The coefficients of the natural logarithm of volume and share turnover are very close to zero and statistically insignificant.

The overall results suggest that the impact of liquidity measures is more strongly felt on the expected return rate of ETFs that tracks major European indexes. As demonstrated above, for every, one cent increase in the spread, the excess return on European ETFs is expected to decrease by approximately 0.021bps or 0.00021% (Panel 3), holding all other independent variables constant.

Consistent with the literature our model also predicts a positive return-illiquidity relationship (see, Amihud (2002)). For the European market, the estimated impact of illiquidity on ETFs excess returns is both economically and statistically significant.

The results of the robustness test are consistent with the results presented above, thus confirming the statistical and economic significance of the liquidity measures, the Fama-French mimic size and value factors, and the market risk premium in determining ETFs excess returns.

The estimations results reported in Panel 3 of **Table 11** shows that the coefficient of the market risk premium, as well as the coefficients of the Fama-French size and value factors, are all statistically significant at 1% level.

<Table 4.6 Robustness Analysis>

The magnitude of the coefficients of spread and spread ratio is even larger for the European sample of ETFs, as reported in panel 3 the GLS coefficients of spread and spread ratio are close to -0.024bp and 1.43bp, -0.00024% and 0.0143% respectively. For the American subsample, no evidence of statistical significance below 5% was found for the coefficients of spread and spread ratio.

The coefficient of illiquidity ranges from 0.441 to 1.423 in the case of the European sample of ETFs, and it is statistically different from zero at 1% level for all regression models (Panels 1 – 3). For the American sample of ETFs, the coefficient of illiquidity exhibits a weaker degree of impact, as it ranges from 0.0322 to 0.0427, and it is only significantly different from zero at a significance level above 1%.

The results for the nonlinear parameters show that the coefficient of the spread variable squared is quite irrelevant economically and statistically for both markets.

The coefficient of the spread ratio squared is positive and only statistically significant for the American subsample of ETFs (Panel 1 and 3), meaning that on average, the excess return of ETFs tracking major benchmark indexes in the U.S. is higher the lower the spread ratio, until a certain point where it will start increasing with the spread ratio.

The estimated turning point (when the expected excess return starts increasing, on average with the spread ratio) for the American subsample of ETFs using robust estimates is given by the spread ratio = $\frac{\hat{\beta}_{spread\ ratio}}{2*\hat{\beta}_{spread\ ratio\ squared}} = \frac{0.00428}{2*0.000144} = 14.86$.

The coefficient of the Amihud illiquidity measure squared is negative and statistically significant at least at 5 % level for the European sample of ETFs, and in the case of American subsample of ETFs, it is statistically significant at 1% level (Panel 2 and 3).

These results are consistent with the literature, as our model predicts a reverse U-shape or concave relationship between ETFs excess returns and the Amihud illiquidity measure. Therefore, the excess returns on ETFs are expected to increase with respect to illiquidity until a certain point where the cost to transact a very illiquid ETF will absorb all the returns. The estimated turning point (where expected excess returns on ETFs starts decreasing with illiquidity, on average) for the European subsample of ETFs using GLS estimates is given by the Amihud Illiquidity measure = $\frac{\hat{\beta}_{Illiquidity}}{2*\hat{\beta}_{Illiquidity\ squared}} = \frac{0.441}{2*28.70} = 0.0077$.

The effect of the interaction term (spread times Amihud illiquidity measure or illiquidity factor) is positive and significantly different from zero at 1% level in the case of the European ETFs, while for the American sample of ETFs it is not statistically significant (Panels 1 – 3).

The coefficients of the interactive term that are statistically significant ranges from 0.111 to 0.238. The impact of spread (illiquidity factor) on the expected excess returns of ETFs that tracks major European indexes are enhanced by changes in the illiquidity factor (spread).

The coefficient of the one-period lag of the dependent variable is negative and statistically significant at 1% level in the case of European sample of ETFs, it ranges from -0.000103 to -0.0144 (Panel 1 and 3). For the American subsample, it is also negative and statistically significant at least at 5% level (Panel 2 and 3).

The dummy variable capturing the effect inherent to the specificity of C40 FP is positive and statistically significant at least at 10% level (Panel 2), while the SPY ETF dummy is not statistically significant. This result means that the ETF tracking the CAC40 indexes is expected to earn on average an excess return of approximately 1bp (0.01%) higher than the remaining ETFs.

In sum, the overall results of the robustness analysis confirm the previous results. We found clear and reliable results concerning the main driven forces of ETFs excess returns, and we demonstrated that the intensity to which the ETFs tracking major European or American benchmark indexes reacts to sudden changes in liquidity tend to diverge.

5. Conclusions

Despite the high popularity of ETFs, studies concerning the risk-return structure of ETFs and its relationship with microstructure factors, such as liquidity are still scarce.

We contribute to the existing literature in two ways, by first providing an in-depth analysis of the performance of the ETFs tracking major European and American benchmark indexes, thus fulfilling the existent gap in the literature focused mainly on the performance of ETFs designed to track major stock market indexes in U.S., while the performance and pricing efficiency of European ETFs are poorly mentioned in the literature.

Second, we provide empirical evidence of the main driven forces of ETFs expected returns rate, which so far, to the best of our knowledge no study has specifically investigated the main determinants of ETFs returns.

Consistent with the literature we found that most ETFs included in our study were unable to beat their underlying benchmark indexes, through the analysis of the Jensen's alpha parameter.

Alternative models to evaluate securities' returns, open to multiple sources of risk rather than one market index factor, seems to be reliable and effective predictors of ETFs returns. We have demonstrated that liquidity measures such as bid-ask spread, spread ratio, and Amihud illiquidity factor are accountable for changes in the expected return rate of ETFs.

We found that our European subsample of ETFs is more sensitive to shocks in liquidity than their American counterparts. The estimations reveal a strong and robust effect of liquidity on the expected excess returns of European equity ETFs.

The expected return rate of European ETFs, as well as their ability to outperform the market, are clearly undermined by sudden changes in liquidity, as they exhibit a highly sensitive response function to shocks in liquidity and inferior performance compared to the ones tracking major benchmark indexes in the U.S.

We concluded that the easiness to which we can sell or buy an asset or the intrinsic cost of engaging in a transaction, do have something to say about ETFs' returns. Hence more complex and highly effective assets pricing models should be able to capture potential sources of risk that are intrinsically related with securities' own market-microstructure.

6. References

Acharya, V. V., and Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics, 77(2), 375–410*. https://doi.org/10.1016/j.jfineco.2004.06.007

Agapova, A. (2011). Conventional mutual index funds versus exchange-traded funds. *Journal of Financial Markets, 14(2), 323–343*. https://doi.org/10.1016/j.finmar.2010.10.005

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5(1), 31–56.* https://doi.org/10.1016/S1386-4181(01)00024-6.

Amihud, Y., and Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17(2), 223–249*. https://doi.org/10.1016/0304-405X(86)90065-6

Authers, J., and Newlands, C. (2016). Exchange traded funds: taking over the markets. *Financial Times*.

Bello, Z. (2012). The investment performance and tracking errors of small-cap ETFs. *Global Journal of Finance and Banking Issues,* 6(6), 12–20.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business, 45(3), 444–455*.

Black, F., Jensen, M., and Scholes, M. (1972). The capital asset pricing model: *Some Empirical Tests. Studies in the Theory of Capital Markets (pp. 79–121).* https://doi.org/10.2139/ssrn.908569

Blitz, D., and Huij, J. (2012). Evaluating the performance of global emerging markets equity exchange-traded funds. *Emerging Markets Review, 13(2), 149–158*. https://doi.org/10.1016/j.ememar.2012.01.004

Blitz, D., Huij, J., and Swinkels, L. (2012). The performance of European index funds and exchange-traded funds. *European Financial Management, 18(4), 649–662*. https://doi.org/10.1111/j.1468-036X.2010.00550.x

Brinson, G. P., Hood, L. R., and Beebower, G. L. (1995). Determinants of portfolio performance. *Financial Analysts Journal, 51(1), 133–138.* https://doi.org/10.2469/faj.v51.n1.1869

Broman, M. S. (2016). Liquidity, style investing and excess comovement of exchange-traded fund returns. *Journal of Financial Markets, 30, 27–53.* https://doi.org/10.1016/j.finmar.2016.05.002

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52(1), 57. https://doi.org/10.2307/2329556.*

Chen, N., Roll, R., and Ross, S. A. (1986). Economic forces and the stock market. *The Journal of Business, 59(3), 383–403.*

Chen, Z., and J. Knez, P. (1996). Portfolio performance measurement: Theory and applications. *The Review of Financial Studies, 9(2), 511–555*.

Cheng, H.-L., Jegadeesh, N., and Wermers, R. (2000). Active mutual fund management: An examination of the stockholdings and trades of fund managers. *The Journal of Financial and Quantitative Analysis, 35(3), 343–368*. https://doi.org/10.2307/2676208

Chordia, T., Roll, R., and Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics, 56(1), 3–28*. https://doi.org/10.1016/S0304-405X(99)00057-4

Cremers, M., Ferreira, M. A., Matos, P., and Starks, L. (2016). Indexing and active fund management: International evidence. *Journal of Financial Economics, 120(3), 539–560*. https://doi.org/10.1016/j.jfineco.2016.02.008

Datar, V. T., Naik, Y. N., and Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets, 1(2), 203–219*. https://doi.org/10.1016/S1386-4181(97)00004-9

De Winne, R., Gresse, C., and Platten, I. (2014). Liquidity and risk sharing benefits from opening an ETF market with liquidity providers: Evidence from the CAC 40 index. *International Review of Financial Analysis, 34, 31–43*. https://doi.org/10.1016/j.irfa.2014.04.003

DeFusco, R. A., Ivanov, S. I., and Karels, G. V. (2011). The exchange traded funds' pricing deviation: Analysis and forecasts. *Journal of Economics and Finance, 35(2), 181–197*. https://doi.org/10.1007/s12197-009-9090-6

Elton, E. J., Gruber, M. J., Comer, G., and Li, K. (2005). Spiders: Where are the bugs? *In Exchange Traded Funds: Structure, Regulation and Application of a New Fund Class (pp. 37–59). Springer, Berlin, Heidelberg*. https://doi.org/10.1007/3-540-27637-8_4

Engle, R., and Sarkar, D. (2006). Premiums-discounts and exchange traded funds. *The Journal of Derivatives, XXXIII (2), 81–87*. https://doi.org/10.1007/s13398-014-0173-7.2

Fama, E. F., and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33(1), 3–56*.

Ferson, W., and Lin, J. (2014). Alpha and performance measurement: The effects of investor disagreement and heterogeneity. *Journal of Finance, 69(4), 1565–1596*. https://doi.org/10.1111/jofi.12165

Flood, C. (2018, April 30). UK and Germany offer widest choice in ETFs and trackers: *Financial Times, pp. 1–2.*

Frino, A., and Gallagher, D. R. (2001). Tracking S&P500 index funds. *Journal of Portfolio Management, 28(1), 44–55*. https://doi.org/10.3905/jpm.2001.319822

Gallagher, D. R., and Segara, R. (2006). The performance and trading characteristics of exchange-traded funds. *Journal of Investment Strategy, 1(2), 49–60*.

Gastineau, G. L. (2004). The benchmark index ETF performance problem. *The Journal of Portfolio Management, 30(2), 96–103*. https://doi.org/10.3905/jpm.2004.319935

Gibbons, M. R., and Ferson, W. (1985). Testing asset pricing models with changing expectations and an unobservable market portfolio. *Journal of Financial Economics, 14(2), 217–236.* https://doi.org/10.1016/0304-405X(85)90015-7

Glode, V. (2011). Why mutual funds "underperform." *Journal of Financial Economics, 99(3), 546–559*. https://doi.org/10.1016/j.jfineco.2010.10.008

Greenwich Associates. (2018). ETFs: Valuable versatility in a newly volatile market. *Greenwich Associates 2017 Exchange-Traded Funds Study.*

Hegde, S. P., and McDermott, J. B. (2004). The market liquidity of DIAMONDS, Q's, and their underlying stocks. *Journal of Banking and Finance, 28(5), 1043–1067*. https://doi.org/10.1016/S0378-4266(03)00043-8

Hilliard, J. (2014). Premiums and discounts in ETFs: An analysis of the arbitrage mechanism in domestic and international funds. *Global Finance Journal, 25(2), 90–107*. https://doi.org/10.1016/j.gfj.2014.06.001

Ibbotson, R. G., Chen, Z., Kim, D. Y. J., and Hu, W. Y. (2013). Liquidity as an investment style. *Financial Analysts Journal, 69(3), 30–44*. https://doi.org/10.2469/faj.v69.n3.4

Idzorek, T. M., Xiong, J. X., and Ibbotson, R. G. (2012). The liquidity style of mutual funds. *Financial Analysts Journal, 68(6), 38–53*. https://doi.org/10.2469/faj.v68.n6.3

Ivanov, S. I. (2013). High- frequency analysis of exchange traded funds' pricing deviation. *Managerial Finance, 39(5).* https://doi.org/10.1108/03074351311313834

Ivanov, S. I. (2016). Analysis of ETF bid-ask spread components. *Quarterly Review of Economics and Finance, 61, 249–259*. https://doi.org/10.1016/j.qref.2016.02.004

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance, 23(2), 389–416*. https://doi.org/10.1111/j.1540-6261.1968.tb00815.x

Kostovetsky, L. (2003). Index mutual funds and exchange-traded funds. *The Journal of Portfolio Management, 29(4), 80–92*. https://doi.org/https://doi.org/10.3905/jpm.2003.319897

L. Treynor, J., and Black, F. (1973). How to use security analysis to improve portfolio selection. *The Journal of Business, 46(1), 66–86.* Retrieved from http://www.jstor.org/stable/2351280

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

Lee, H. C., Tseng, Y. C., and Yang, C. J. (2014). Commonality in liquidity, liquidity distribution, and financial crisis: Evidence from country ETFs. *Pacific Basin Finance Journal, 29, 35–58*. https://doi.org/10.1016/j.pacfin.2014.03.006

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance, 20(4), 338–342*. https://doi.org/https://doi.org/10.1111/j.1540-6261.1965.tb02930.x

Malkiel, B. G. (2003). Passive investment strategies and efficient markets. *European Financial Management, 9(1), 1–10*. https://doi.org/10.1111/1468-036X.00205

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, (2), 55–75*. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Marshall, B. R., Nguyen, N. H., and Visaltanachoti, N. (2018). Do liquidity proxies measure liquidity accurately in ETFs? *Journal of International Financial Markets, Institutions and Money*. *Elsevier Ltd*. https://doi.org/10.1016/j.intfin.2018.02.011

Meziani, A. S. (2016). Exchange-traded funds: Investment practices and tactical approaches *(1st ed.), (pp. 1–395).* Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-39095-0

Milonas, N. T., and Rompotis, G. G. (2006). Investigating European ETFs: The case of the Swiss exchange traded funds. *The Annual Conference of HFAA, 1–28*.

Pástor, Ľ., Stambaugh, R. F., and Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics, 116(1), 23–45*. https://doi.org/10.1016/j.jfineco.2014.11.008

Peltomäki, J. (2017). Beta as a determinant of investor activity in sector exchange-traded funds. *Quarterly Review of Economics and Finance, 65, 137–145*. https://doi.org/10.1016/j.qref.2016.06.006

Pope, P. F., and Yadav, P. K. (1994). Discovering errors in tracking error. *The Journal of Portfolio Management, 20(2), 27 LP-32*. Retrieved from http://jpm.iijournals.com/content/20/2/27.abstract

Poterba, B. J. M., and Shoven, J. B. (2016). American economic association exchange-traded funds: A New Investment Option for Taxable Investors. *The American Economic Review, Vol. 92, No. 2, Papers and Proceedings of the One Hundred F, 92(2).*

Roll, R. (1992). A mean/variance analysis of tracking error. *The Journal of Portfolio Management, 18(4), 13–22*. https://doi.org/10.3905/jpm.1992.701922

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13(3), 341–360*. https://doi.org/10.1016/0022-0531(76)90046-6

Sharpe, W. F. (1964). Capital asset prices: A theory of equilibrium under conditions of risk. *The Journal of Finance, 19(3), 425-442*. https://doi.org/10.2307/2977928

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business, 39(S1), 119*. https://doi.org/10.1086/294846

Shin, S., and Soydemir, G. (2010). Exchange-traded funds, persistence in tracking errors and information dissemination. *Journal of Multinational Financial Management, 20(4–5), 214–234*. https://doi.org/10.1016/j.mulfin.2010.07.005

Sortino, F. A., and Price, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing, 3(3), 59 LP-64.* Retrieved from http://joi.iijournals.com/content/3/3/59.abstract

Stafford, P. (2016). The differences between US and European ETF markets. *Financial Times*.

Statman, M. (1987). How many stocks make a diversified portfolio? *The Journal of Financial and Quantitative Analysis, 22(3), 353–363*. Retrieved from http://www.jstor.org/stable/2330969

Subrahmanyam, A. (1991). A theory of trading in stock index futures. *The Review of Financial Studies, 4(1), 17–51*. https://doi.org/https://doi.org/10.1093/rfs/4.1.17

Svetina, M. (2010). Exchange traded funds: Performance and competition. *Journal of Applied Finance, 20(November), 130–145.* https://doi.org/10.2139/ssrn.1303643

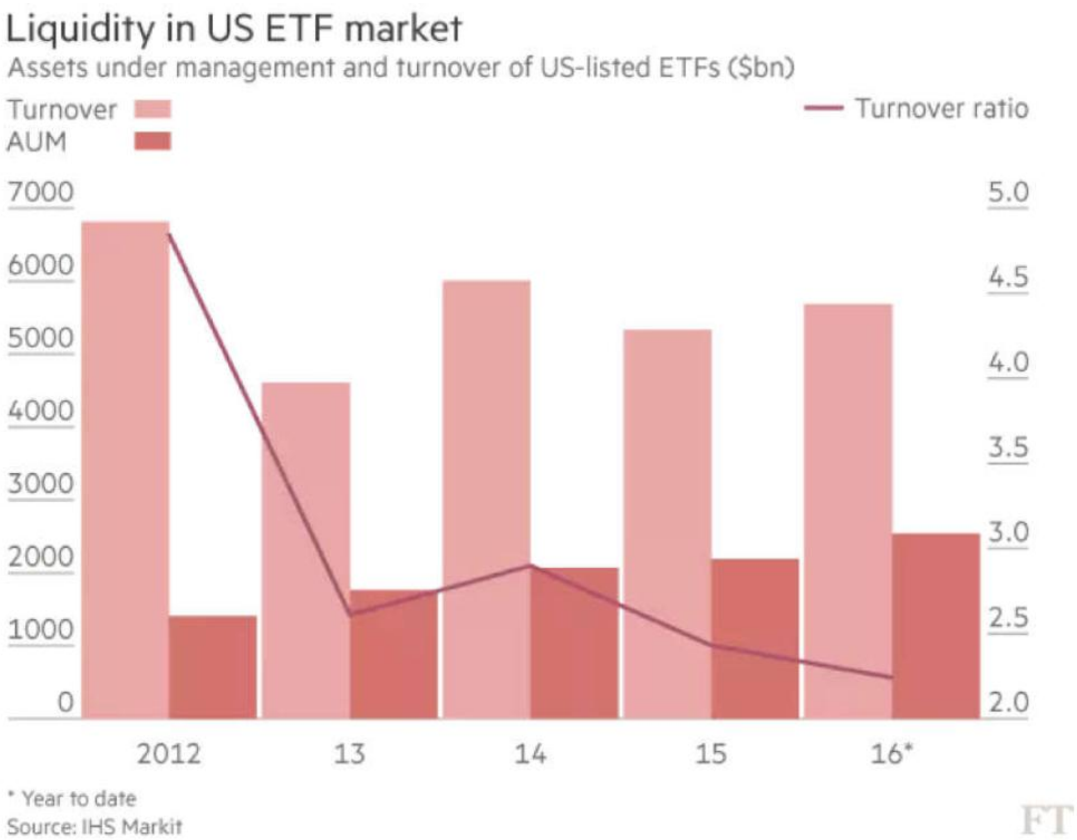
Thirumalai, R. S. (2003). Active vs. passive ETFs. *Kelley School of Business Working Paper. Indiana University, Bloomington.*

Yiannaki, S. M. (2015). ETFs performance Europe - A good start or not? *Procedia Economics and Finance, 30(15), 955–966.* https://doi.org/10.1016/S2212-5671(15)01346-5

Zvie, B., Kane, A., and Alex J., M. (2013). *Investments (10th ed.).* McGraw-Hill Education.

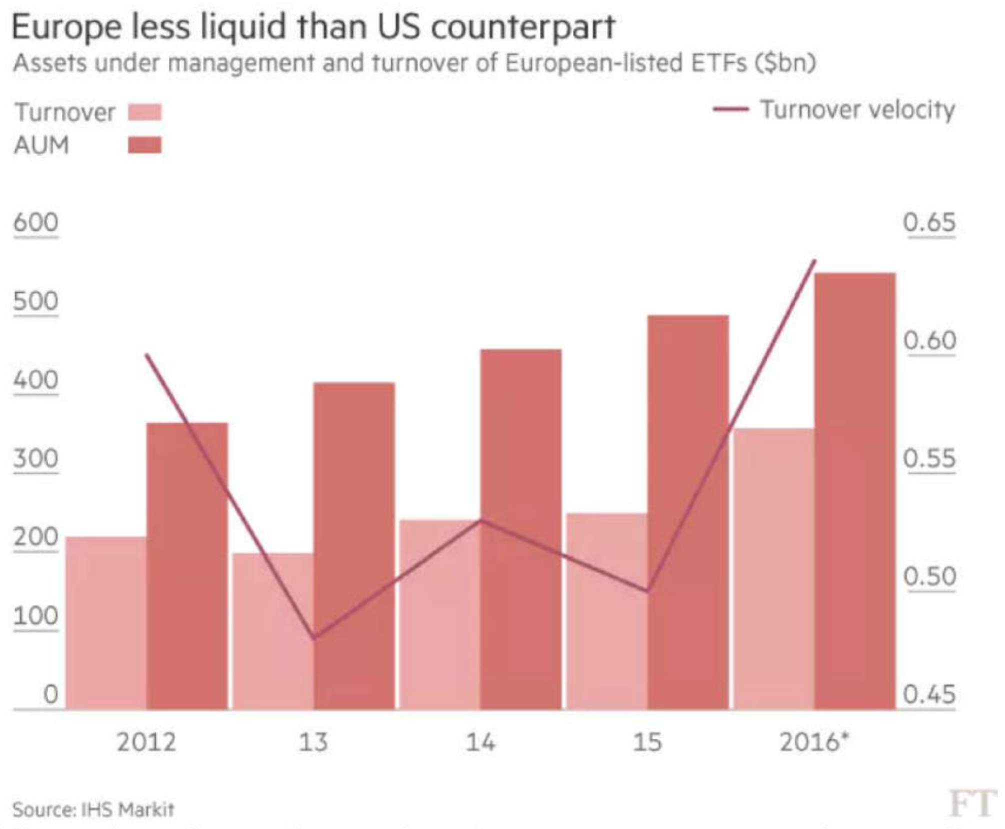
7. List of Figures

Figure 1.1 Liquidity in U.S. ETF Market



Source: Financial Times, 2016 (www.ft.com)

Figure 1.2 Liquidity in the Europe ETF Market



Source: Financial Times, 2016 (www.ft.com)

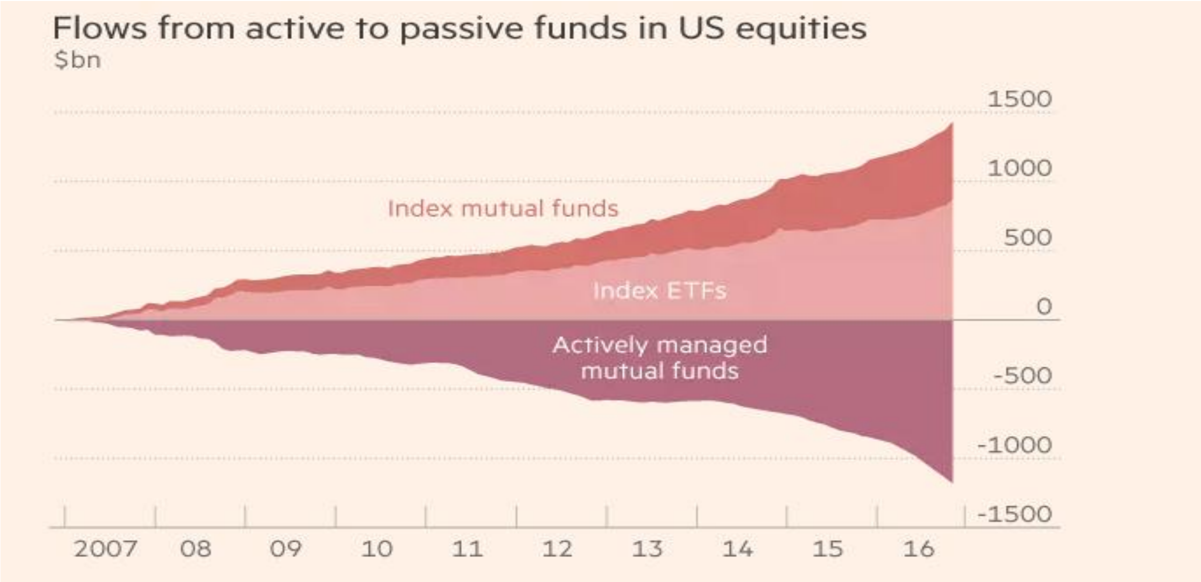


Figure 2.1 ETFs Take Over

Source: Financial Times, 2016 (www.ft.com)

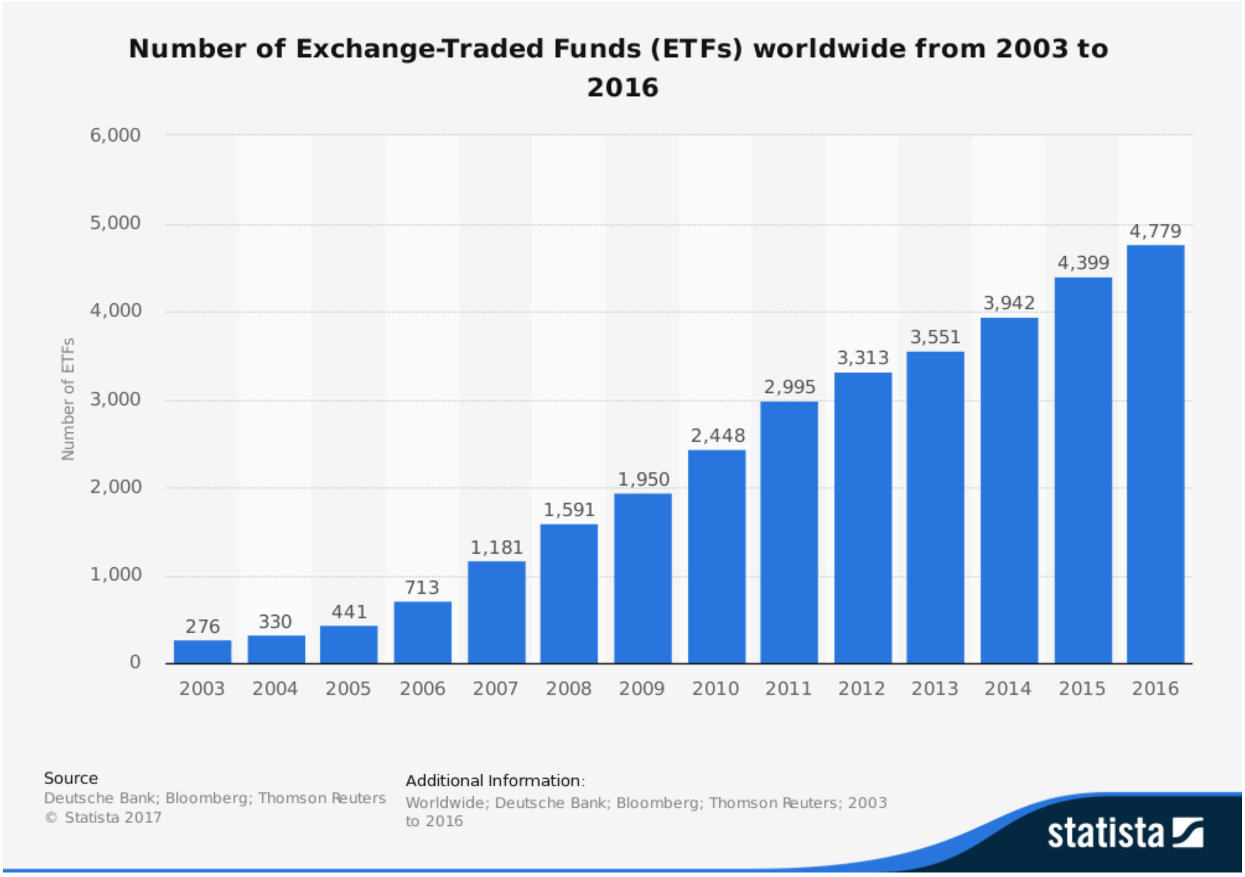
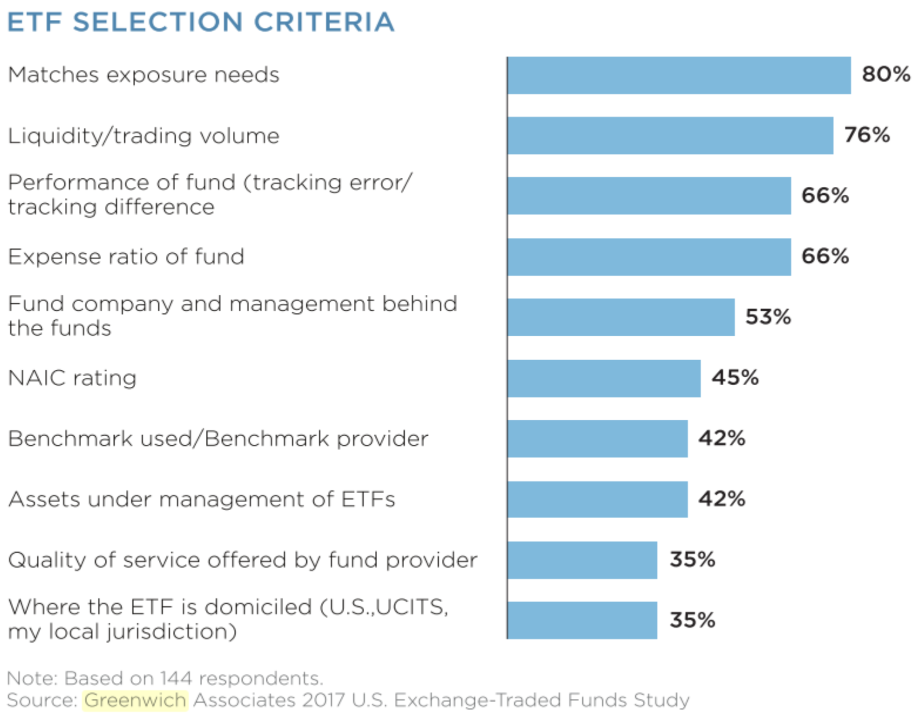


Figure 2.2 Evolution of ETFs Products Worldwide

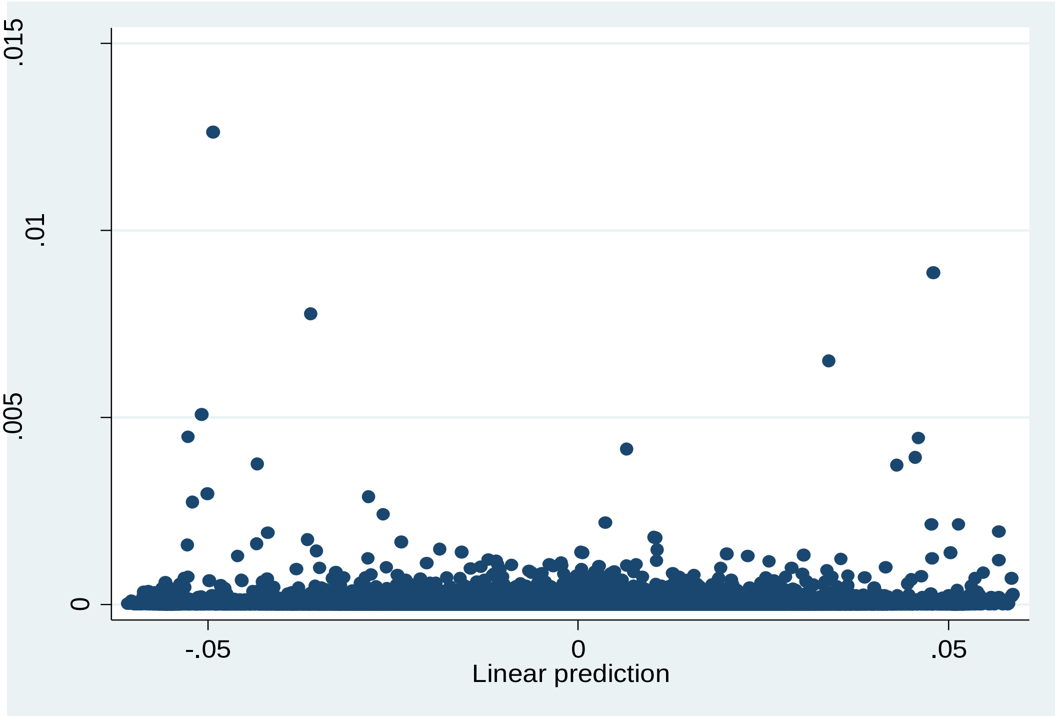
Source: Statista, 2018 (www.statista.com)

Figure 2.3 ETFs Selection Criteria



Source: Greenwich Associates, 2017

Figure 3.1 Plot of Squared Residuals and Linear Prediction



8. List of Tables

Table 1 Summary Statistics of the overall sample and description of variables

This table provides quick summary statistics of the entire sample and describes the variables used. The values in parenthesis correspond to the standard deviation.

Variables	Description	Units	Minimum	Maximum	Mean (Std. deviation)	Source
Ri	Return on ETF.	Percentage	-3.773	3.482	0.023 (1.190)	Bloomberg
Rm	Return on the benchmark.	Percentage	-3.716	3.480	0.025 (1.189)	Bloomberg
Ri	Excess return on ETF.	Percentage	-5.648	5.222	0.040 (1.960)	Bloomberg
Rm	Excess return on market indexes.	Percentage	-5.642	5.199	0.043 (1.961)	Bloomberg
Spread	Closing bid-ask spreads.	In euros	0.000	6.223	0.331 (0.940)	Bloomberg
Spread ratio	Closing bid-ask spreads as a percentage of ETFs prices.	Percentage	0.000	9.874	0.579 (1.524)	Bloomberg
ILLIQ	Amihud (2002) illiquidity Factor.	Percentage	-0.819	0.895	0.00148 (0.080)	Bloomberg
Volume	Number of shares traded.	Number of shares	248	253,000,000.00	15,900,000.00 (44,100,000.00)	Bloomberg
Turnover	Daily trading volume divided by the end of day shares outstanding.	Number of shares	68.672	464,000.00	44,600.00 (85,300.00)	Bloomberg
P/B ratio	Is the price-to-book ratio.	Ratio	0.005	94.098	2.281 (4.543)	Bloomberg
MKT	Current market capitalization.	In millions of euros	27.096	219,839.437	34,856.277 (38,059.052)	Bloomberg
NAV	Net asset Value.	In euros	4.011	226.235	107.827 (48.947)	Bloomberg
HML	Mimic value factor.	Percentage	-1.322	1.354	0.001 (0.460)	Bloomberg

SMB	Mimic size factor	Percentage	-1.375	1.313	-0.003 (0.420)	Bloomberg

Table 2 Funds' Characteristics

This table presents the descriptive Statistics of fund's characteristics. For each ETF we report the computed mean (rounded to units). Volume stands for the average number of shares traded by each ETF; Spread ($) is the closing bid-ask spread; P/B ratio is the price-to-book ratio; MKT is the current market capitalization in millions of euros; NAV is the net asset value expressed in euros; Turnover is the daily trading volume divided by the end of day shares outstanding, and N stands for the number of observations. For all variables, the mean is calculated based on daily observations.

ETF	Benchmark	Volume	Spread	P/B ratio	MKT	NAV	Turnover	N
SPY	S&P 500	153,000,000	0.020	0.729	91,058.346	119.617	224,000	3025
QQQ	Nasdaq	77,300,000	0.004	0.864	22,096.241	54.065	186,000	3021
DIA	DJIA	9,320,000	0.091	3.535	7,780.478	106.489	126,000	3025
IWN	iShares Russell 2000	1,620,000	0.729	0.257	3,523.676	61.296	28,845	3025
VTV	CRSP U.S. Large Cap Value Index	1,050,000	0.714	0.412	7,290.973	50.805	4,976	1221
VTI	CRSP US Total Market Index	2,480,000	0.064	0.790	22,446.012	60.957	5,816	1567
IJH	S&P Midcap 400 Index	966,000	0.851	0.482	11,153.803	81.106	9,088	3021
IVV	S&P500	3,570,000	0.583	0.774	31,638.256	120.072	16,663	3021
VOO	S&P500	1,270,000	0.673	0.784	11,621.047	74.154	10,935	1701
DAXEX	DAX 30	177,000	0.066	0.617	7,099.959	69.574	3,360	2965
C40 FP	CAC 40	58,193	0.119	0.430	593.797	49.179	5,235	2999
EWI	Italy MISC 25/50	578,000	0.524	0.312	332.854	28.191	21,654	2736
PPP	PSI 20	64,794	0.041	0.384	23.994	3.415	11,037	1193
ISF LN	FTSE 100	6,300,000	0.012	1.291	3,118.725	7.447	13,421	2915
LYXIB	IBEX	69,703	0.946	0.439	253.427	97.220	30,230	2606
Bell 20	BEL 20 index	4,765	0.346	0.286	40.149	40.511	4,499	2971
TDT NA	AEX Index	35,572	0.141	0.711	44.412	22.610	20,055	1556
SMICH A	Swiss Market Index	49,539	0.096	0.302	670.851	57.209	5,002	2921
Sum of N	45,489							
ETFs	18							

Table 3 Model 1 Descriptive Statistics

This Table provides the descriptive statistics of the main variables included in our first model testing the abnormal return. We compute the mean and the standard deviation (value in parenthesis) for each variable. ri stands for the daily return; rm is the market return; Ri and Rm are the expected excess return earned by the ETF and by the inherent benchmark, respectively, and NAV return is the return resulting from daily changes in the net asset value of each ETF. N is the number of observations.

ETF	Benchmark	ri	Ri	rm	Rm	NAV return	N
SPY	S&P 500	0.0210	0.0347	0.0218	0.0365	0.0214	3030
		(1.240)	(1.903)	(1.256)	(1.904)	(1.254)	
QQQ	Nasdaq	0.0404	0.0532	0.0417	0.0543	0.0414	3020
		(1.333)	(1.982)	(1.368)	(1.988)	(1.368)	
DIA	DJIA	0.0212	0.0354	0.0321	0.0467	0.0214	3030
		(1.187)	(1.908)	(1.183)	(1.903)	(1.181)	
IWN	iShares Russell 2000	0.0212	0.0359	0.0217	0.0344	0.0216	3030
		(1.601)	(2.008)	(1.619)	(2.00260)	(1.619)	
VTV	CRSP U.S. Large Cap Value Index	0.0489 (0.954)	0.0236 (1.906)	0.0598 (0.950)	0.0344 (1.898)	0.0488 (0.956)	1220
VTI	CRSP US Total Market Index	0.0452 (1.0131)	0.0702 (2.0239)	0.0533 (1.0155)	0.0777 (2.0202)	0.0452 (1.0188)	1570
IJH	S&P Midcap 400 Index	0.0308	0.04420	0.0312	0.0452	0.0310	3020
		(1.381)	(1.936)	(1.396)	(1.933)	(1.394)	
IVV	S&P500	0.0216	0.0367	0.0210	0.0365	0.0209	3020
		(1.239)	(1.908)	(1.256)	(1.905)	(1.256)	
VOO	S&P500	0.0481	0.0584	0.0483	0.0581	0.0482	1700
		(0.9792)	(2.0299)	(0.981)	(2.0238)	(0.982)	
DAXEX	DAX 30	0.0281	0.0576	0.0274	0.0579	0.0240	2970
		(1.336)	(1.898)	(1.324)	(1.918)	(1.322)	
C40 FP	CAC 40	0.0250	0.0421	0.0145	0.0329	0.0240	3000
		(1.381)	(1.938)	(1.384)	(1.929)	(1.384)	
EWI	Italy MISC 25/50	-0.00210	0.0241	0.00540	0.0344	-0.00530	2740
		(1.682)	(1.923)	(1.542)	(1.965)	(1.534)	
PPP	PSI 20	-0.0186	0.0436	-0.00810	0.0528	-0.00390	1190
		(1.300)	(2.149)	(1.298)	(2.133)	(1.299)	
ISF LN	FTSE 100	0.00940	0.0213	0.0107	0.0227	0.0102	2920
		(1.257)	(1.955)	(1.254)	(1.950)	(1.255)	

LYXIB	IBEX	0.00180	0.0216	0.0200	0.0391	-0.000600	2610
		(1.564)	(2.0828)	(1.565)	(2.0692)	(1.570)	
Bell 20	BEL 20 index	0.0237	0.0428	0.0146	0.0335	0.0260	2970
		(1.143)	(1.922)	(1.222)	(1.932)	(1.150)	
TDT NA	AEX Index	0.0171	0.0275	0.0167	0.0269	0.0141	1560
		(1.090)	(1.975)	(1.103)	(1.972)	(1.134)	
SMICHA	Swiss Market Index	0.0272	0.0489	0.0337	0.0578	0.0256	2920
		(1.029)	(1.989)	(0.994)	(1.983)	(2.088)	
Sum of N	45,520						
ETFs	18						

Table 4 Model 2 Descriptive Statistics

This table presents the summary statistics of the main variables used in the second model aimed to test the impact of liquidity in determining ETFs returns. Here ETFs are sort by market and the descriptive statistics is then presented in two panels. Panel A reports the statistics of ETFs tracking major European indexes and panel B reports the statistics for ETFs that track major U.S. market indexes. Ri and Rm correspond to the expected excess return on the ETF and the market, respectively; Spread is the closing bid-ask spreads; Spread ratio is the closing bid-ask spreads as a percentage of ETFs' prices; ILLIQ is the Amihud, (2002) illiquidity factor; HML and SML are the mimics size and value factors developed in Fama and French (1993); Log Vol and Log Turn are the natural logarithm of ETFs volume and turnover respectively.

Variable	Mean	Std. Deviation	Min	Max	Skewness	Kurtosis	Median	N
Panel A Europe								
Ri	0.037	1.969	-5.648	5.222	-13.940	349.875	0.070	22862
Rm	0.040	1.972	-5.642	5.199	-14.828	348.197	0.073	22862
HML	-0.005	0.522	-1.322	1.354	3.401	332.417	-0.006	22862
SMB	-0.004	0.548	-1.375	1.313	-9.297	321.074	0.004	22862
Spread	0.265	0.784	-	6.223	5.192	33.130	0.050	22862
Spread ratio	0.584	1.470	0.000	9.874	416.587	2202.512	0.114	22862
ILLIQ	0.001	0.094	-0.819	0.895	-14.190	819.851	0.004	22862
Log Vol	10.695	2.626	5.513	18.298	0.330	2.660	10.418	22862
Log Turn	8.212	1.720	4.229	13.048	-0.348	2.568	8.479	22862
Panel B U.S.								
Ri	0.043	1.952	-5.648	5.222	-19.628	359.769	0.097	22627
Rm	0.045	1.949	-5.642	5.199	-19.575	359.942	0.101	22627
HML	0.006	0.389	-1.322	1.354	-9.979	449.123	0.015	22627
SMB	-0.003	0.229	-1.375	1.313	33.079	723.448	-0.007	22627
Spread	0.398	1.070	-	6.223	3.858	18.334	0.031	22627
Spread ratio	0.573	1.577	0.000	9.874	399.954	2028.161	0.033	22627
ILLIQ	0.002	0.064	-0.701	0.611	-27.776	1084.877	0.003	22627
Log Vol	15.427	1.988	9.136	19.348	0.447	2.222	14.895	22627
Log Turn	10.254	1.500	7.158	13.048	0.171	1.815	10.088	22627

Table 5 Correlation Matrix

This table presents the pairwise correlations among the variables included in our study.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Ri	1.000								
(2) Rm	0.982*	1.000							
(3) HML	-0.026*	-0.044*	1.000						
(4) SMB	0.020*	0.027*	0.085*	1.000					
(5) Spread	-0.000	0.000	-0.001	-0.000	1.000				
(6) Spread ratio	0.001	0.001	-0.000	-0.003	0.929*	1.000			
(7) ILLIQ	0.246*	0.215*	-0.155*	0.127*	-0.010	-0.015*	1.000		
(8) Log Vol	-0.002	-0.001	0.011	0.002	-0.092*	-0.132*	-0.013*	1.000	
(9) Log Turn	-0.004	-0.003	0.009	0.000	-0.018*	-0.031*	-0.026*	0.809*	1.000

* shows significance at the 1% level

Table 6 Reward-to-volatility Indicators (Tabulated by ETF)

This table presents the computed indicators of reward-to-volatility for each ETF, using Sharpe, Treynor, and Sortino ratios. TE1 correspond to the absolute tracking error and TE2 is the standardized version of tracking error.

ETF	Sharpe ratio (%)	Treynor ratio (%)	Sortino ratio (%)	TE1	TE2	N
SPY	1.824	0.0348	2.649	-0.000007	0.00174	3,025
QQQ	2.686	0.0536	3.988	-0.000013	0.00149	3,021
DIA	1.854	0.0352	2.698	-0.000109	0.00167	3,025
IWN	1.787	0.0361	2.646	-0.000005	0.00218	3,025
VTV	1.239	0.0236	1.812	-0.000109	0.00096	1,221
VTI	3.470	0.0702	5.223	-0.000080	0.00080	1,567
IJH	2.280	0.0443	3.337	-0.000004	0.00160	3,021
IVV	1.922	0.0367	2.796	0.000007	0.00141	3,021
VOO	2.877	0.0583	4.360	-0.000002	0.00090	1,701
DAXEX	3.032	0.0611	4.532	0.000007	0.00641	2,965
C40 FP	2.174	0.0423	3.246	0.000105	0.00262	2,999
EWI	1.250	0.0289	1.902	-0.000075	0.01139	2,736
PPP	2.031	0.0440	3.024	-0.000104	0.00335	1,193
ISF LN	1.092	0.0214	1.626	-0.000013	0.00187	2,915
LYXIB	1.037	0.0218	1.602	-0.000182	0.00419	2,606
Bell 20	2.227	0.0438	3.310	0.000091	0.00373	2,971
TDT NA	1.391	0.0276	2.106	0.000004	0.00183	1,556
SMICHA	2.457	0.0490	3.743	-0.000065	0.00348	2,921
Sum of N	45,598					
ETFs	19					

Table 7 Reward-to-Volatility (Tabulated by Market)

This table presents the indicators of reward-volatility for the Europe and the U.S. sub samples of ETFs, using Sharpe, Treynor, and Sortino ratios. TE1 correspond to the absolute tracking error

Variable	Sharpe ratio (%)	Treynor ratio (%)	Sortino ratio (%)	TE1	TE2	N
Europe	1.880	0.038	2.829	-0.0000182	0.00533	22862
U.S.	2.182	0.043	3.211	-0.0000290	0.00157	22627
Sum of N	45,598					
ETFs	18					

and TE2 is standardized version of tracking error.

Table 8 Model 1 Testing Abnormal return (Tabulated by Market)

This table provides the regression results of the first model testing the abnormal return. The dependent variables are the excess return on ETFs (Ri), and the independent variables are the excess return on the market (Rm). Additionally, we provide results for the mean comparison test. The computed difference between ETFs and benchmark returns, as well as the estimated t-stat are presented on the right-side of the table. ETFs are sorted by market, the Europe and the U.S. subsamples of ETFs are composed by ETFs that tracks the performance of major European and U.S. market stock indexes, respectively.

	(1)		(2)	
Model	OLS robust		Mean Difference Test	
Market	Europe	U.S.	Europe	U.S.
Variables	Ri	Ri	Ri-Rm	Ri-Rm
Rm	0.965***	0.998***		
	(0.00225)	(0.000489)		
Constant	-0.0000122	-0.0000282***		
	(0.00003350)	(0.00000981)		
R-Squared	0.934	0.994		
Difference in returns			-.0000261	-.0000289
t-stat			-0.142	-0.158
N	22862	22627	22862	22627
ETFS	9	9	9	9

[*] $p < 0.1$, [**] $p < 0.05$, [***] $p < 0.01$ indicate statistical significance at 10, 5, and 1 percent levels. Values in parenthesis correspond to the robust standard errors.

Table 9 Model 1 Testing abnormal return (tabulated by ETF)

This table provides the regression results of the first model testing the abnormal return based on OLS. The dependent variables are the excess return on ETFs (Ri), and the independent variables are the excess return on the market (Rm).

ETF	Benchmark	Constant	Rm	N	R-Squared
SPY	S&P 500	-0.00000625 (0.0000318)	0.997*** (0.00258)	3025	0.993
QQQ	Nasdaq	-0.00000867 (0.0000270)	0.993*** (0.00172)	3021	0.995
DIA	DJIA	-0.000111*** (-0.0000305)	1.005*** (0.00320)	3025	0.993
IWN	iShares Russell 2000	-0.00000257 (0.0000397)	0.995*** (0.00239)	3025	0.990
VTV	CRSP U.S. Large Cap Value Index	-0.000109*** (0.0000276)	1.002*** (0.00141)	1221	0.998
VTI	CRSP US Total Market Index	- 0.0000814*** (0.0000203)	1.001*** (0.00111)	1567	0.999
IJH	S&P Midcap 400 Index	-0.00000266 (0.0000293)	0.997*** (0.00189)	3021	0.994
IVV	S&P500	0.00000706 (0.0000258)	1.000*** (0.00178)	3021	0.995
VOO	S&P500	-0.00000311 (0.0000221)	1.002*** (0.00122)	1701	0.998
DAXEX	DAX 30	0.0000286 (0.000116)	0.942*** (0,00852)	2981	0.902
C40 FP	CAC 40	0.0000947* (0.0000483)	0.997*** (0.00252)	3021	0.984
EWI	Italy MISC 25/50	-0.0000315 (0.000209)	0.833*** (0.0181)	2758	0.710
PPP	PSI 20	-0.000101 (0.0000971)	0.993*** (0.00458)	1193	0.977
ISF LN	FTSE 100	-0.0000121 (0.0000379)	0.998*** (0.00195)	2935	0.990
LYXIB	IBEX	-0.000180** (0.0000814)	0.991*** (0.00507)	2616	0.965
Bell 20	BEL 20 index	0.0000982 (0.0000681)	0.977*** (0.00459)	2973	0.966
TDT NA	AEX Index	0.00000150 (0.0000464)	0.996*** (0.00207)	1557	0.992
SMICHA	Swiss Market Index	-0.00000606 (0.0000647)	0.998*** (0.00390)	2937	0.973
Sum of N	45,598				
ETFs	19				

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate statistical significance at 10, 5, and 1 percent levels. Values in parenthesis correspond to the robust standard errors.

Table 10 Determinants of ETFs returns

This table provides the regression results of the second model. We use random effects models, generalized least squares models, and robust OLS to assess the impact of liquidity on ETFs linked to major U.S. and European indexes markets. The dependent variables are the expected excess return on ETFs (Ri) and the independent variables are: The return on the market (Rm); mimic value and size factors (HML and SMB); closing bid-ask spreads (Spread); the closing bid-ask spreads as a percentage of ETFs' prices (Spread ratio); ILLIQ is the Amihud, (2002) illiquidity factor; HML and SML are the mimics size and value factors developed in Fama and French (1993); Log Vol and Log Turn are the natural logarithm of ETFs volume and turnover, respectively.

	(1)		(2)		(3)	
Model	OLS Robust		RE		GLS	
Market	Europe	U.S.	Europe	U.S.	Europe	U.S.
Variables	Ri	Ri	Ri	Ri	Ri	Ri
Rm	0.955***	0.999***	0.953***	0.999***	0.987***	0.999***
	(0.00292)	(0.000652)	(0.00166)	(0.000516)	(0.00104)	(0.000394)
HML	0.152***	0.0240*	0.147***	0.000674	0.0295***	-0.0102***
	(0.0136)	(0.00824)	(0.00659)	(0.00392)	(0.00435)	(0.00317)
SMB	-0.0857***	-0.0400***	-0.0964***	-0.0608***	-0.00989***	-0.0581***
	(0.0115)	(0.0153)	(0.00628)	(0.00656)	(0.00400)	(0.00537)
Spread	0.0000164	0.000000162	-0.000188**	0.00000976	-0.000207***	0.0000370
	(0.0000184)	(0.00000892)	(0.0000913)	(0.0000373)	(0.0000672)	(0.0000261)
Spread Ratio	-0.00115	-0.00000492	0.00930*	-0.000260	0.00835**	-0.00206
	(0.000930)	(0.000693)	(0.00485)	(0.00254)	(0.00391)	(0.00181)
ILLIQ	1.562***	0.0433	1.453***	-0.00841	0.406***	0.00673
	(0.0875)	(0.0433)	(0.0353)	(0.0167)	(0.0235)	(0.0135)
Log Vol	-0.00000123	0.00000521	-0.00000444	0.00000421	-0.00000559	-0.00000140
	(0.0000127)	(0.0000116)	(0.0000159)	(0.0000107)	(0.00000703)	(0.00000662)
Log Turn	0.00000777	-0.00000114	0.0000111	-0.0000109	-0.00000902	0.0000000691
	(0.0000251)	(0.0000156)	(0.0000245)	(0.0000138)	(0.0000141)	(0.00000778)
Constant	-0.0000644	0.00000485	-0.0000711	0.0000143	-0.0000163	-0.0000209
	(0.000157)	(0.0000937)	(0.000163)	(0.0000827)	(0.0000946)	(0.0000536)
R-Squared	0.941	0.994	0.939	0.994		
N	22862	22627	22862	22627	22862	22627
ETFs	9	9	9	9	9	9

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate statistical significance at 10, 5, and 1 percent levels. Values in parenthesis correspond to the robust standard errors.

Table 11 Robustness Analysis

This table provides the regression results of the second model, after controlling for ETF dummies, non-linear, and interactions variables. We use random effects models, generalized least squares models, and robust OLS to assess the impact of liquidity on ETFs linked to major U.S. and European indexes markets. The dependent variables are the expected excess return on ETFs (Ri) and the independent variables are: the return on the market (Rm); mimic value and size factors (HML and SMB); closing bid-ask spreads, absolute and squared values (Spread and Spread^2); iSpread is the closing bid-ask spreads (Spread) multiplied by the Amihud, (2002) illiquidity factor (ILLIQ); Spread as percentage of ETFs' prices, absolute and squared values (Spread ratio and Spread ratio^2); absolute and squared value of the Amihud, (2002) illiquidity factor (ILLIQ and ILLIQ^2); Log Vol and Log Turn are the natural logarithm of ETFs volume and turnover, respectively; L1 Ri is one period lag of the dependent variable; we include two ETFs dummies as control variables to account for the distinctive characteristics of the CAC 40 and S&P 500 market indexes.

Model	(1) OLS Robust		(2) GLS		(3) RE	
Market	Europe	U.S.	Europe	U.S.	Europe	U.S.
Variables	Ri	Ri	Ri	Ri	Ri	Ri
Rm	0.953***	0.998***	0.986***	0.999***	0.953***	0.998***
	(0.00286)	(0.000513)	(0.00120)	(0.000446)	(0.00189)	(0.000565)
HML	0.146***	-0.00651	0.0367***	-0.0133***	0.146***	-0.00651
	(0.0122)	(0.00611)	(0.00503)	(0.00371)	(0.00762)	(0.00443)
SMB	-0.101***	-0.0846***	-0.0125***	-0.0753***	-0.101***	-0.0846***
	(0.0110)	(0.0124)	(0.00462)	(0.00630)	(0.00727)	(0.00746)
Spread	-0.000199	0.0000540	-0.000241***	0.0000547*	-0.000199*	0.0000540
	(0.000200)	(0.0000535)	(0.0000816)	(0.0000316)	(0.000106)	(0.0000423)
Spread^2	-0.0000000830	-0.0000000231**	-0.0000000644	-0.0000000104	-0.0000000830	-0.0000000231**
	(0.0000000972)	(0.00000000903)	(0.0000000404)	(0.00000000928)	(0.0000000557)	(0.0000000114)
iSpread	0.238***	0.0147	0.111***	-0.00119	0.238***	0.0147
	(0.0842)	(0.0215)	(0.0292)	(0.0113)	(0.0401)	(0.0132)
Spread ratio	0.0108	-0.00428	0.0143***	-0.00380*	0.0108*	-0.00428
	(0.0127)	(0.00401)	(0.00482)	(0.00219)	(0.00565)	(0.00289)
Spread ratio^2	0.0000745	0.000144***	-0.00000790	0.0000632	0.0000745	0.000144***
	(0.0000776)	(0.0000555)	(0.0000900)	(0.0000565)	(0.0000633)	(0.0000688)
ILLIQ	1.423***	0.0427	0.441***	0.0322*	1.423***	0.0427**
	(0.0927)	(0.0352)	(0.0298)	(0.0167)	(0.0449)	(0.0202)
ILLIQ^2	-37.43	-7.476	-28.70**	-40.65***	-37.43**	-7.476
	(59.52)	(28.01)	(12.87)	(9.004)	(18.04)	(9.774)
Log Vol	-0.00000861	0.00000567	-0.00000709	-0.0000107	-0.00000861	0.00000567
	(0.0000141)	(0.0000145)	(0.00000859)	(0.00000901)	(0.0000185)	(0.0000140)
Log Turn	0.0000205	-0.0000307**	0.0000246	-0.00000349	0.0000205	-0.0000307*
	(0.0000274)	(0.0000154)	(0.0000175)	(0.00000973)	(0.0000287)	(0.0000160)
"C40 FP"	0.0000997		0.0000925*		0.0000997	
	(0.0000762)		(0.0000542)		(0.000111)	
"SPY"		0.0000408		0.0000416		0.0000408
		(0.0000464)		(0.0000339)		(0.0000426)
L1_Ri	-0.0144***	-0.00101	-0.000103	-0.000930**	-0.0144***	-0.00101*
	(0.00267)	(0.000655)	(0.00114)	(0.000426)	(0.00183)	(0.000531)
Constant	-0.0000107	0.000164	-0.000139	0.000145*	-0.000107	0.000164
	(0.000194)	(0.000130)	(0.000124)	(0.0000795)	(0.000195)	(0.000166)

R-Squared	0.941	0.995			0.941	0.995
N	17315	17500	17315	17500	17315	17500
ETFs	9	9	9	9	9	9

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate statistical significance at 10, 5, and 1 percent levels. Values in parenthesis correspond to the robust standard errors.

9. Appendix

Figure 4.1 Regression Fitted Line by ETF

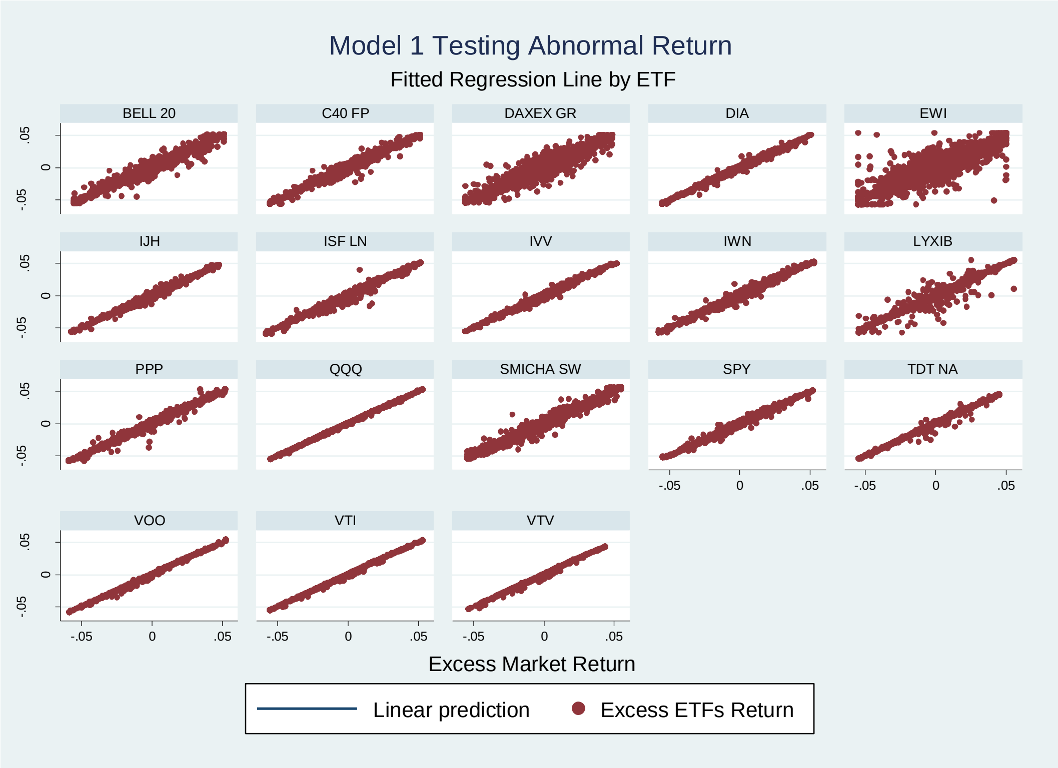


Figure 4.2 Regression Fitted Line by Market

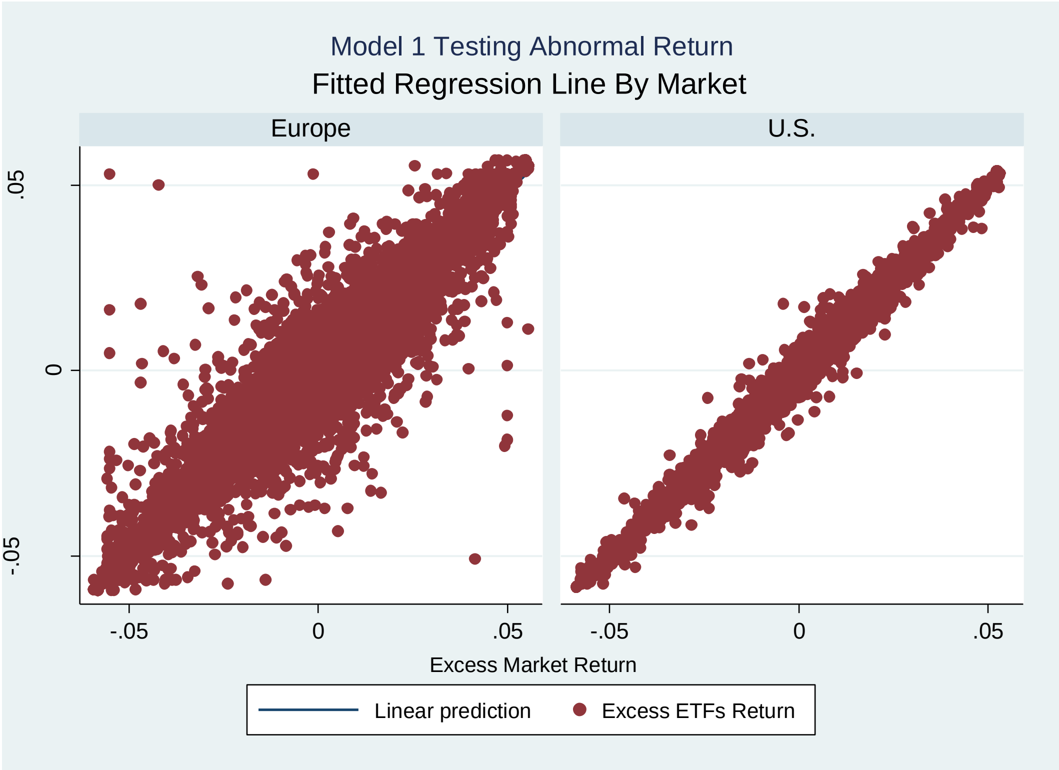


Table 12 Estimations Results of the fixed effect model

This table provides the regression results of the second model. We use fixed effect models to assess the impact of liquidity on ETFs linked to major U.S. and European indexes markets. The dependent variables are the expected excess return on ETFs (Ri) and the independent variables are: The return on the market (Rm); mimic value and size factors (HML and SMB); closing bid-ask spreads (Spread); the closing bid-ask spreads as a percentage of ETFs' prices (Spread ratio); ILLIQ is the Amihud, (2002) illiquidity factor; HML and SML are the mimics size and value factors developed in Fama and French (1993); Log Vol and Log Turn are the natural logarithm of ETFs volume and turnover, respectively.

	(2)	
Model	FE	
Market	Europe	U.S.
Variables	Ri	Ri
Rm	0.953***	0.999***
	(0.0247)	(0.000689)
HML	0.147	0.000679
	(0.105)	(0.0213)
SMB	-0.0964	-0.0605
	(0.0737)	(0.0441)
Spread	-0.000196	-0.00000274
	(0.000162)	(0.0000185)
Spread Ratio	0.0108*	0.0000768
	(0.00552)	(0.00103)
ILLIQ	1.454*	-0.0105
	(0.760)	(0.0411
Log Vol	-0.0000150	-0.0000171
	(0.0000136)	(0.0000114)
Log Turn	0.0000543	-0.0000141
	(0.0000331)	(0.000144)
Constant	-0.000319*	0.000380**
	(0.000157)	(0.000137)
R-Squared	0.940	0.994
N	22862	22627
ETFs	9	9

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate statistical significance at 10, 5, and 1 percent levels. Values in parenthesis correspond to the robust standard errors.