LAPPEENRANTA UNIVERSITY OF TECHNOLOGY
School of Business and Management
Strategic Finance and Business Analytics

Joonas Timonen

ETF PRICING INEFFICIENCIES – EVIDENCE FROM LONDON STOCK EXCHANGE

Master's Thesis

2019

1st Examiner: Professor Eero Pätäri

2nd Examiner: Associate Professor Sheraz Ahmed

ABSTRACT

Author:	Joonas Timonen
Title:	ETF pricing inefficiencies – evidence from London Stock Exchange
Faculty:	LUT School of Business and Management
Master's program:	Strategic Finance and Business Analytics
Year:	2019
Master's thesis:	82 pages, 13 tables, 5 figures and 6 appendices
Examiners:	Professor Eero Pätäri
	Associate Professor Sheraz Ahmed
Keywords:	ETF, NAV, pricing inefficiency, ETF arbitrage

ETFs (Exchange-traded funds) have quickly established themselves as important investment vehicles in modern day finance. The aim of this thesis is to study deviations between ETFs' trading prices and their net asset values (NAVs), the phenomenon also known as ETF mispricing. This thesis applies methodologies of previous ETF research to find out the current state of ETF pricing efficiency in London Stock Exchange (LSE). Based on the obtained results it is evident that ETF mispricing also occurs in LSE. Despite the fact that, on average, ETFs are trading very close to NAV, ETFs can exhibit very high and persistent premiums from time to time. Magnitude of these premiums is related to specific barriers of ETF arbitrage. Among these barriers are ETF-specific bid-ask spread, volatility of underlying assets and time-zone differences between ETF trading market and market of underlying assets. Also, overall market sentiment is found to influence mispricing, as mispricing increases simultaneously with FTSE 100 VIX index. For uninformed investors, these findings represent potential pitfalls that could erase cost advantages offered by ETFs. For informed investors, knowledge of mispricing offers opportunities for additional profits. Analysis on long-short trading based on mispricing provides further evidence of the existence of the phenomenon. ETF trading prices exhibit predictable patterns, which to some extent could be utilized to enhance potential profits of ETF investors. However, these arbitrage profits are highly sensitive to transaction costs, especially to bid-ask spread and thus, they may not be systematically exploitable if actual trading costs are considered.

TIIVISTELMÄ

Tekijä:	Joonas Timonen
Aihe:	ETF-rahastojen hinnoittelun tehottomuus – todisteita Lontoon pörssistä
Tiedekunta:	LUT School of Business and Management
Pääaine:	Strategic Finance and Business Analytics
Vuosi:	2019
Pro Gradu:	82 sivua, 13 taulukkoa, 5 kuviota, 6 liitettä
Tarkastajat:	Professori Eero Pätäri
	Apulaisprofessori Sheraz Ahmed
Hakusanat:	ETF, NAV, hinnoittelun tehottomuus, ETF-arbitraasi

ETF-rahastot ovat nopeasti nousseet tärkeäksi sijoitusvälineeksi nykypäivän rahoitusmaailmassa. Tutkielman tarkoitus on selvittää eroa ETF-rahastojen kaupankäyntihinnan ja nettoarvon (NAV) välillä, eli ETF-rahastojen hinnoitteluvirhettä. Tutkielmassa käytetään aiemman ETF-tutkimuksen metodologioita selvittämään ETF-rahastojen tämänhetkistä hinnoittelutehokkuutta Lontoon pörssissä. Saatujen tulosten perusteella on selvää, että ETF-rahastojen hinnoitteluvirheitä esiintyy myös Lontoon pörssissä. Huolimatta siitä, että keskimäärin ETF-rahastoilla käydään kauppaa lähellä nettoarvoa, ajoittaisia voimakkaita ja pysyviä preemioita voi esiintyä. Preemioiden voimakkuus liittyy esteisiin toteuttaa ETF-arbitraasia. Näiden esteiden joukossa ovat ETF-kohtaiset osto- ja myyntihinnan ero, kohde-etuuksien volatiliteetti sekä aikaero ETF-rahaston kaupankäyntipörssin ja kohdemarkkinan välillä. Myös yleisen markkinasentimentin voidaan osoittaa vaikuttavan hinnoitteluvirheeseen, sillä hinnoitteluvirhe kasvaa samanaikaisesti FTSE 100 VIX indeksin kanssa. Ilmiöön perehtymättömälle sijoittajalle löydökset tarkoittavat mahdollisia vaaroja, jotka voivat nopeasti poistaa ETF-rahastojen tarjoamat hintaedut. Ilmiöön perehtyneelle sijoittajalle hinnoitteluvirhe tarjoaa mahdollisuuksia lisätuottoihin. Analyysi hinnoitteluvirhettä hyödyntävästä pitkää ja lyhyttä kauppaa käyvästä strategiasta tarjoaa lisätodisteita ilmiön olemassaolosta. ETF-rahastojen kaupankäyntihinnat osoittavat ennakoitavia liikeitä, joita voidaan osittain hyödyntää lisätuottojen mahdollistamiseksi ETF-sijoittajille. Kuitenkin nämä arbitraasituotot ovat herkkiä kaupankäyntikustannuksille, erityisesti osto- ja myyntihinnan erolle, eivätkä välttämättä ole systemaattisesti hyödynnettävissä todelliset kaupankäyntikustannukset huomioiden.

ACKNOWLEDGEMENTS

As this thesis is now completed I'm grateful to all who helped me during this process.

I would like to thank Professor Eero Pätäri for your comments and guidance. Also, I want to express my gratitude towards Professor Mikael Collan, Associate Professor Sheraz Ahmed and other staff of LUT for providing opportunity and environment to develop myself. Thanks also to all friends at LUT. You made the time spent at Lappeenranta memorable.

Finally, I would like to thank my family for continuous support during my studies. Especially thanks to Belinda for your encouragement and support during these years.

Espoo 31th of January 2019

Joonas Timonen

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ETF	Exchange-traded fund
NAV	Net asset value
iNAV	Intraday NAV
LSE	London Stock Exchange
EMH	Efficient market hypothesis
AP	Authorized participant (Performer of ETF arbitrage)
CAPM	Capital asset pricing model

1. INTRODUCTION

During the recent decade Exchange-traded funds (ETFs) have quickly emerged as a popular alternative for more traditional investments. ETFs are funds that are listed and traded vía exchanges and are tracking the performance of selected indices, stocks, commodities, bonds or baskets of securities. Offering various benefits, they are currently considered a viable challenger to traditional mutual funds. Due to the recent popularity ETF industry has been able to obtain increasing amount of capital invested. Raising popularity and capital invested makes these investment vehicles important subject in modern day finance.

Regarding capital invested in ETFs, Forbes (2017) reports that ETF industry has grown already into over 4 trillion-dollar business in year 2017. Growth is remarkable as most of currently listed ETFs have their inception dates only after the year 2000. Having developed in US, ETFs are today expanding globally, and more funds are being listed in also smaller exchanges. Figure 1 demonstrates growth of the ETF industry globally. Values are obtained from ETF survey published by Ernst and Young (2017). Estimates for year 2020 indicate staggering growth as ETF industry is expected to double in coming years.

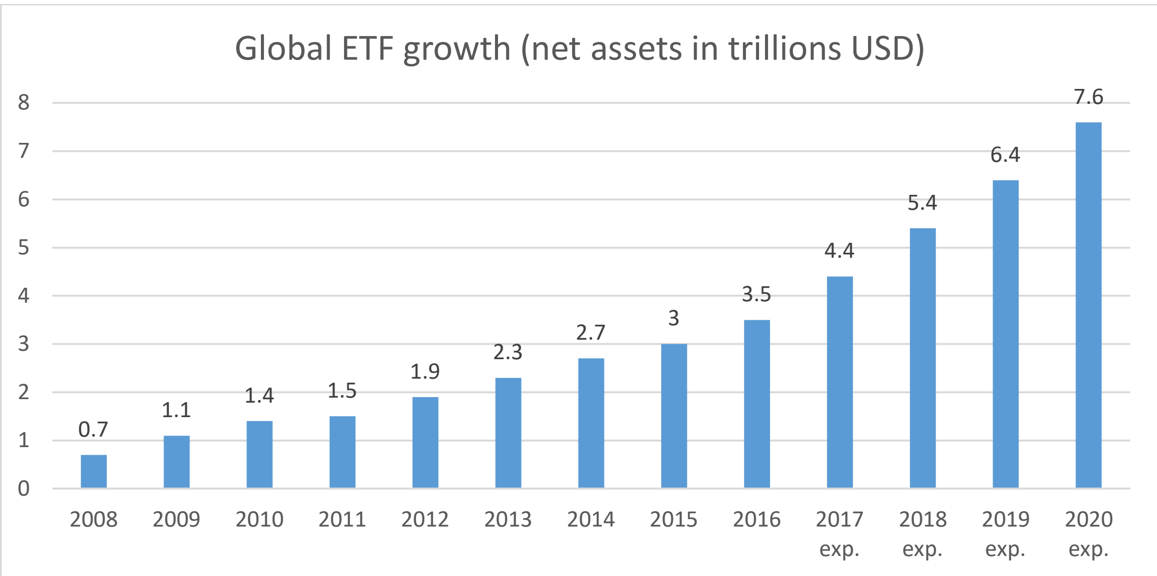


Figure 1: Global growth of ETF industry (Ernst and Young, 2017)

Exponential growth is a result of several benefits that are attractive for investors. One of the benefits of ETFs for investor is cost efficiency. As aim of ETF is to track selected index or assets, they do not require active portfolio management. Since ETFs are rather passive instruments, managing ETF requires less active management from fund managers, especially compared to active mutual funds that aim to beat their benchmark indices. Often passive management translates to lower expenses for ETF investors as fees are lower for ETFs compared to active mutual funds. Regarding other benefits, ETFs can offer broad exposure to diversified indices with ease. Acquiring even one ETF share allows this broad exposure which otherwise might be difficult for an average investor to obtain. Also, from this perspective cost efficiency can be considerable. For example, more exotic assets or markets are available for even smaller investors vía ETFs for reasonable costs. Lastly, flexibility compared to mutual funds is an issue that is usually discussed with ETFs. As ETFs are exchange traded, investors can buy or sell ETF shares vía stock exchange in similar manner as traditional shares. This type of flexibility also allows intraday trading, which is not possible in conventional mutual fund investments. Due to these benefits ETFs have gained popularity as an alternative for more traditional investment vehicles such as stocks or mutual funds.

Exponential growth rate of ETF industry and popularity among investors are features that have caught attention of academic research. Following the growing number of ETFs, also increasing amount of research has been conducted to discover how ETFs function. In addition to the previously-mentioned benefits associated with ETFs, academic research has found some possible disadvantages. A phenomenon known as ETF pricing inefficiency or ETF mispricing is one of such issues. Recent academic research has found that some ETFs exhibit mispricing when fund's trading price is examined against fund's net asset value (NAV). NAV can be argued to reflect the fundamental value of ETFs as ETFs are tracking performance of other assets. When the trading price of ETF is above NAV fund is said to trade at premium. Vice versa when trading price is below NAV ETF is said to trade at discount or negative premium (for clarity and consistency with previous literature both cases are expressed as term premium from here on even if premium is negative).

ETF mispricing is not a new phenomenon as it has also been documented even at early stages of ETF industry. The early studies of Elton et al (2002) and Engle and Sarkar (2006) examined the efficiency of ETF pricing in the US. On average, ETFs tend to trade quite efficiently. However, larger deviations occur from time to time and especially with certain asset classes underlying ETFs. ETFs following international or illiquid assets tend to exhibit larger premiums compared to other alternatives (Hilliard, 2014). Pricing efficiency becomes even more important issue in the future as ETF industry attracts more capital and more complicated and exotic ETFs are incepted to be traded in various markets. Petäjistö (2017) gives an estimate for current economic magnitude of ETF mispricing: $20 billion yearly solely in the US.

Since academia has found mispricing to exist many types of research on ETF mispricing and premiums have been implemented. Studies regarding ETF pricing efficiency can involve testing mispricing with trading strategy to see whether economically and statistically significant excess profits can be acquired. Moreover, these types of studies are mainly considered from the perspective of informed investor trying to capitalize on mispricing. Past studies ranging from Jares and Lavin (2004) to Petäjistö (2017) have found that mispricing of ETFs has offered arbitrage opportunities.

An other approach used in ETF research has been examinations of factors contributing to mispricing. Studies in this area aim to explain what drives ETF trading prices further away from NAVs. Notable studies include Delcoure and Zhong (2007) and Ackert and Tian (2008). ETF premiums are magnified by barriers of ETF arbitrage, which is a process used by institutional investors to ensure efficient pricing. ETFs having larger barriers to conduct this arbitrage are more prone to mispricing as market participants are more unable to arbitrage the occurred mispricing away. Later studies of Petäjistö (2017) and Ben-David et al. (2012) have linked premiums to overall market factors. When overall uncertainty increases and the availability of arbitrage capital vía capital markets is scarce, ETFs tend to trade with more inefficient prices.

To this date most research on ETFs is done using the US domiciled ETFs. Since the ETF industry grows rapidly and more ETFs are listed in frontier stock exchanges, it is valuable to extend literature on ETF pricing efficiency to unresearched markets. Generally, a limited amount of research regarding ETFs trading outside the US has been published. Theories and models developed for the US-based ETFs can be tested to see whether similar dynamics in ETF pricing can also be found in other market. Also, studies which consider multiple ETF classes are relatively scarce, as academic research has mostly focused on equity ETFs.

Due to the relatively narrow research regarding ETFs trading outside the US, there are clear research gaps regarding ETFs trading in London Stock Exchange (LSE). Despite being one of the major exchanges in Europe, studies involving ETFs trading at LSE are scarce. Chelley and Park (2011) and Mohamad et al. (2015) represent studies using LSE traded ETFs. However, they focus on different research questions than ETF mispricing as Chelley and Park (2011) examined intraday trading characteristics of ETFs, while Mohamad et al. (2015) discussed relation of ETFs and shorting in LSE.

1.1 Objectives and research questions

Number of ETFs in LSE is growing with similar rapid manner than ETFs globally. Today over 2000 ETFs are traded in LSE making it one of major hubs for European ETF markets. The purpose of this thesis is to discover current state of ETF mispricing and possible determinants for the phenomenon in LSE. The overall research problem of this thesis can be answered by following research questions:

Q1: Is there evidence of ETF mispricing in LSE?

Past ETF research has found significant deviations between ETF trading price and its NAV in several different markets. Since ETFs are tracking the performance of another index or asset, its fundamental value should be approximately fund's NAV. Because both ETFs and underlying securities are publicly traded, ETFs should have a trading price that is close to

NAV if investors are rational. Despite this reasoning there has been evidence of both economically and statistically significant premiums. Evidence also suggests premiums to be persistent even for number of days indicating consistent inefficiency in pricing of some ETFs. The purpose is to examine whether ETFs traded in LSE are subject to this kind of phenomenon.

Q2: What determinants contribute to magnitude of ETF premiums and mispricing in LSE?

ETFs have unique arbitrage mechanism to establish boundaries for premiums. If deviation between ETF trading price and NAV increases large enough it is arbitraged away by selected institutional investors. Moreover, arbitrage process pushes prices towards equilibrium. Since arbitrage carries costs, small premiums are likely to occur. However, there are constraints for ETF arbitrage that might magnify premiums. These constraints can be related to market conditions or ETF characteristics. Since most theories and models are developed focusing on the US-based ETFs it is reasonable to test whether similar factors can explain magnitude of premiums and mispricing in LSE.

Q3: Is daily ETF mispricing significant enough to be exploited with trading strategies?

ETF mispricing can present opportunity to make excess profits. Relatively simple trading strategies based on exploiting deviation between ETFs trading price and NAV are found to produce considerable profits even after transaction costs. From this perspective active trading strategy can be used as complementary analysis for overall mispricing. If ETFs in LSE were efficiently priced, any attempt to profit from mispricing should not yield abnormal returns.

Q4: Does mispricing and its determinant vary between different types of ETFs?

Equity ETFs have attracted most capital and popularity among ETFs. However, there are also other types of ETFs investing to other types of securities, for example, bonds or commodities. Despite number of different ETF categories available there is relatively scarce amount of literature available on other categories than equities. Studies involving multiple ETF categories indicate that there are differences in pricing efficiency across the categories. Therefore, it is reasonable to examine possible differences also in LSE.

To answer the above-presented research questions this thesis adopts the methodology used in previous ETF research. To identify mispricing, deviations between ETF trading price and NAV are estimated and tested with statistical hypotheses. To see whether mispricing can be exploited in trading, a long-short trading strategy is employed. Determinant factors for premium magnitude and mispricing are tested with different types of regression analysis. Both market level and ETF level factors are used to explain the phenomenon. The sample data consists of 201 ETFs with daily data for a time period of 2.1.2014 to 29.12.2017. Several types of ETFs based on their Morningstar style categories are included in the analysis to see whether there are differences between their categories. The used approach follows closely Petäjistö (2017), who also examined cross-sectional ETF mispricing, ETF trading strategy and premium determinants in combined analysis.

Main contribution of this thesis is the extension of ETF literature to a new market. ETF literature outside the US is scarce and to the best of my knowledge, similar mispricing studies have not been done in LSE. The second contribution is a cross-sectional analysis which is done by including different ETF categories in the sample. Only handful of previous studies report the results for this kind of mixture of different ETFs. Third, the sample of ETFs features relatively recent data. Since the expansion of ETF industry is exponential, the use of recent data allows to perceive whether ETFs have developed in terms of efficiency. Lastly this thesis is among the first studies trying to adjust for stale NAVs in the case of European ETFs. The issue of stale NAVs is considered important for ETFs having trading hours and underlying assets in different time-zones. To this date the stale NAV adjustment has only been applied in studies using the US-based ETF samples.

1.2 Structure

This thesis is structured as follows. First, general theoretical background of pricing efficiency and ETF pricing mechanism are discussed in section 2, which also includes literature review on previous studies on ETF mispricing. The next section describes in detail how the employed data is analyzed. It also introduces the employed models, as well as the reasoning for their usage in hypotheses testing. Section 4 presents empirical results. The final section contains conclusions, limitations of the results and suggestions for further research.

2. THEORETICAL FRAMEWORK

2.1 Efficient market hypothesis

Efficient market hypothesis (EMH) presents general theoretical framework for pricing of publicly traded assets. Pricing of ETFs should also follow EMH as ETFs and their underlying assets are both publicly traded. The concept of EMH was first presented by Fama (1970) and is since widely studied subject in the field of finance. According to EMH, price of an asset should reflect available information. Moreover, statement argues that all relevant information regarding the asset should be incorporated in its price. Therefore, investors should be able to trade assets at their fair price. Trading at fair price restricts possibilities to trade under- or overpriced assets further restricting possibilities to make abnormal returns or "beat the market".

To maintain pricing efficiency, asset prices should react correctly to any new information arriving to the market. EMH also assumes investors to act rationally when making investment decisions. Elton (2003, 431) points out that these should be treated as two separate concepts: informational efficiency and market rationality. However, both are key concepts for EMH to hold. If asset prices adjust correctly to the arriving information and investors act rationally when trading assets, EMH should hold. (Fama 1970)

When assets are priced correctly and they react to new information, investors should expect fair return for their investment. If EMH holds investors would not be able to benefit by using any kind of information in their decision-making as relevant information should be already priced. In such situation investors cannot make excess returns by utilizing different investment strategies. Simply buying and holding asset would be the best investment strategy. If other strategy is found profitable, for example arbitrage opportunities arise, rational investors would replicate this found strategy until profitability disappears. If EMH holds, the only way to obtain higher returns is to select riskier assets into investment portfolio. (Fama 1970)

Modern finance recognizes three forms of market efficiency: weak, semi strong and strong. Bodie et al. (2005, 370-373) explain these as follows:

Weak-form efficiency: Asset prices reflect all past information published. This, for example, contains historical prices. In weak-form efficient markets investors are not able to use this kind of information to make excess profits compared to the buy-and-hold strategy. It can be stated that prices follow random walk and are not predictable. If markets are weak-form efficient technical analysis becomes useless. However, even weak-form efficiency does not suggest that assets are priced at equilibrium all the time. Moreover, it states that strategies using past information are not able to provide constant excess profits.

Semi-strong-form efficiency: This form expands the coverage from weak-form efficiency. In addition to all past information, also all publicly available information is priced according to the semi-strong efficiency form. Any attempt to exploit fundamental analysis cannot provide excess returns. For example, investors cannot benefit from the use of company-specific data such as annual reports in their investment strategies.

Strong-form efficiency: The widest form of EMH. It states that all possible information is incorporated into asset prices. In addition to all previously presented information, it also covers private information. When strong-form efficiency holds insiders cannot use private information to generate excess returns. Strong-form efficiency can be considered as rather theoretical concept as private information is not available for public.

In real world asset prices do not always act according to EMH. Vast amount of empirical research on EMH suggests that violations against theory and different forms of it do exist. Examples of research against EMH include exploiting different market anomalies, possibilities to conduct arbitrage and behavioral finance. Each viewpoint represents potential challenges for proponents of EMH. Considering evidence found regarding EMH it is reasonable to conclude that EMH does not hold all the time. (Ross et al. 2013, 452-457).

However, despite the evidence against EMH it is usable framework as it sets rational reasoning behind asset pricing. This reasoning is applicable for publicly traded securities such as stocks or ETFs alike.

2.2 ETF pricing and arbitrage process

ETFs exhibit unique characteristics compared to more traditional investment vehicles. ETFs are investment vehicles tracking performance of index or another asset. By purchasing one ETF share investor gains exposure to all assets underlying the ETF. Since ETF is tracking another assets, the fundamental or fair value of ETF is approximately value of assets underlying it. Moreover, NAV of ETF should reflect the price which ETF trades on the secondary market.

According to London Stock Exchange (2014) and large ETF issuer Vanguard (2013) there are primarily two methods how ETFs aim to replicate performance of their underlying assets: physical and synthetic. In physical method ETF straightforward owns assets it aims to track. Another way to establish ETF is synthetic tracking. Here ETF does not physically own the assets. Instead, derivatives are used to track the assets of interest. According Vanguard (2013), a swap contract is a common way to establish ETF with synthetic replication method.

Regardless of the way ETF is established the concept of EMH should apply. If investors traded rationally and price adjustment to new information occurred as stated in EMH, the price of ETF should not deviate much from its NAV. This stems from the fact that a rational investor would not pay unnecessary premium for ETF because underlying assets could also be bought instead. If deviation occurs, ETF can be either under- or overpriced against NAV. In academia, significant deviations between ETF trading price and NAV are known as ETF mispricing or ETF pricing inefficiency. If such a pricing inefficiency occurs, ETFs have unique arbitrage process to help setting prices back to equilibrium (Delcoure and Zhong 2007). This process is also known as ETF creation/redemption process. Figure 2 from London Stock Exchange (2014) demonstrates how ETF shares are created or redeemed and how they are then listed to be traded by market participants. This mapping is closely linked to the ETF arbitrage aiming to ensure efficient pricing. The creation/redemption process has remained the same since early documentations of Gastineau (2001) on the first generation of ETFs.

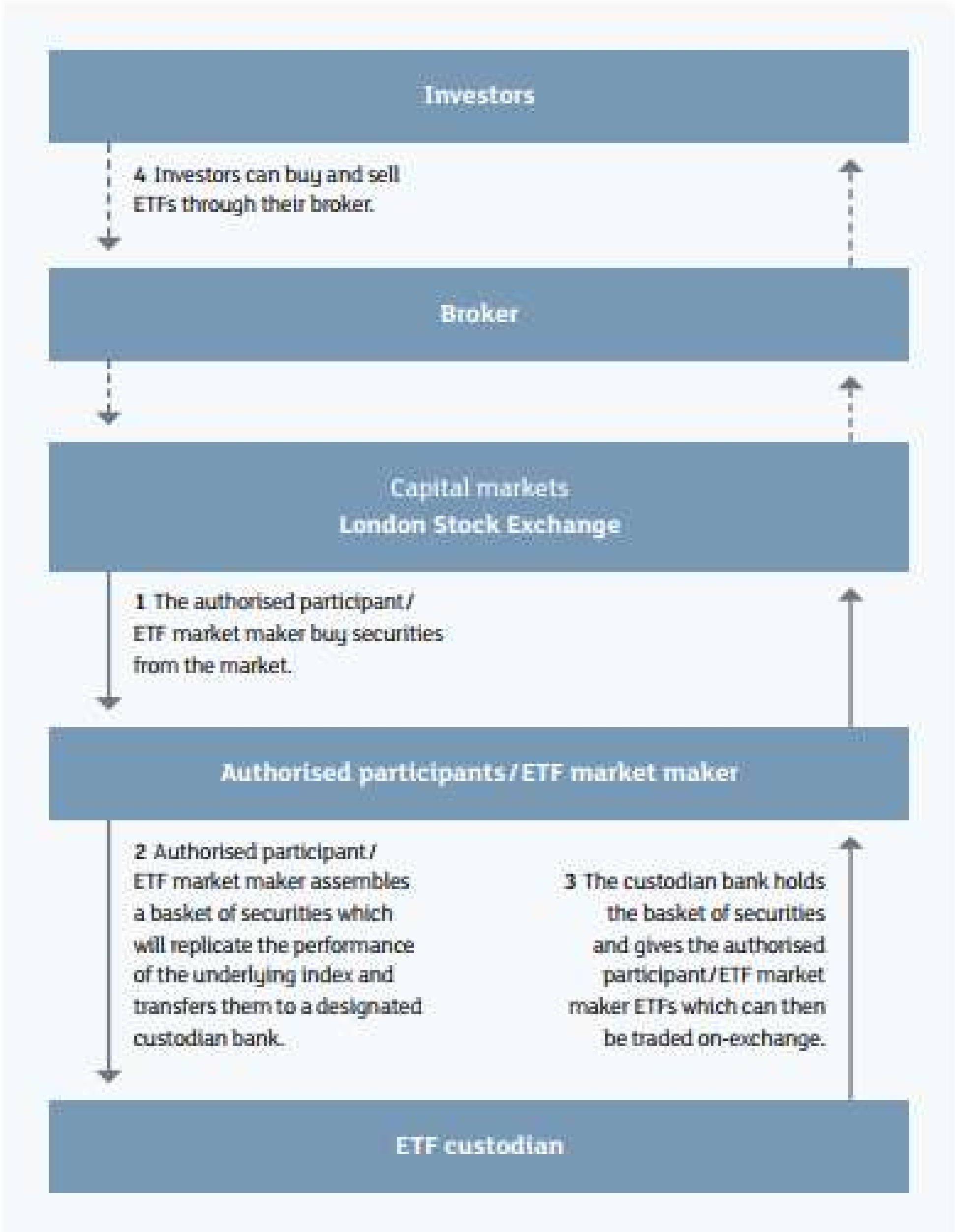


Figure 2: ETF creation/redemption process (London Stock Exchange, 2014)

The process for creating or redeeming ETF shares starts with an authorized participant (AP) and an issuer of ETF. When ETF share is issued APs act as market makers and liquidity providers. APs are institutional entities which have been granted AP status. London Stock

Exchange (2017) provides list of 21 APs acting in LSE. Notable APs are, for example, Deutche Bank, Morgan Stanley and UBS. The creation and redemption of ETF shares requires participation of both ETF issuer and AP. Each day ETF issuers publishes a creation basket – a list of securities which they are willing to accept in exchange for creating ETF shares. AP willing to acquire new ETF shares delivers securities to the ETF issuer in exchange for new ETF shares. Creation and redemption occurs in creation or redemption units which typically consist of larger quantities of ETF shares. For example, number of 50 000 shares is reported by Vanguard (2013). Also, process may involve some fees to be paid by AP. If AP wants to convert ETF shares back into actual assets similar process occurs but vice versa. Moreover, ETF issuer also offers a redemption basket which is a list of securities they are willing to offer in exchange for ETF shares. Here AP delivers ETF shares to ETF issuer in exchange for these securities. In case of synthetic ETFs, Vanguard (2013) informs the process to be similar but instead of delivery of physical assets, cash transactions take place. The described redemption and creation processes between ETF issuer and AP is known as primary market for ETF shares. (Blackrock 2017; Vanguard 2013; Deville 2008).

The secondary market for ETF shares occurs when APs introduce created ETF shares to be traded in public vía stock exchange. When AP has acquired ETF shares it has possibility to trade them in exchange in similar manner than regular shares. These are the shares that other market participants can buy and sell vía their brokers. Thus, secondary markets accommodate majority of the ETF traders. However, due to their involvement both in primary and secondary markets, APs can manage the amount of ETF shares available for public trading. (Vanguard 2013; Blackrock 2017)

Due to their ability to operate in both primary and secondary markets, APs are key players regarding ETF pricing efficiency by their involvement in ETF arbitrage. ETF arbitrage is argued to ensure efficient pricing in the secondary markets. ETF arbitrage takes place when APs utilize their opportunity to create and redeem ETF shares to make arbitrage profit. If AP spots ETF trading at premium against its NAV it can simply buy assets underlying ETF, convert them to ETF shares and finally sell obtained shares at the secondary markets to gain arbitrage profit. The process is similar but reverse when AP spots discount at ETF price against its NAV. Here AP would buy ETF shares from secondary market and convert them to assets underlying ETF with ETF issuer to make arbitrage profits. These arbitrage actions by APs should establish boundaries on ETF pricing in secondary markets. Moreover, they should ensure that ETF's secondary market price tracks fund's NAV closely. If large deviations occur, they are arbitraged away by APs. However, there is no obligation for APs to perform arbitrage activities despite them noticing price deviations. Thus, APs can choose to act only when arbitrage profit is significant enough to be capitalized. For this reason, small deviation between ETF and NAV are likely to occur. Since there are costs and uncertainty involved in the arbitrage process, APs might not consider small deviations profitable enough to be arbitraged. If arbitrage action does not occur, deviations may also persist. (Blackrock 2017; Petäjistö 2017)

From theoretical perspective, if ETF arbitrage mechanism functions properly, it should eliminate any significant pricing inefficiencies quickly ensuring fair ETF pricing for secondary market participants. In practice, arbitrage mechanism may not function as efficiently as the theory suggests. Empirical evidence of ETF mispricing has occurred across different ETF markets. Results obtained by academic studies indicate barriers in ETF arbitrage process. These potential barriers create difficulties and uncertainty for APs to conduct the previously-described ETF arbitrage. Moreover, they decrease APs motivation to engage to the arbitrage process and magnify premiums. In such scenario ETFs can trade at significant and persistent premiums in secondary market as there is no arbitrage activity to push prices back towards NAVs.

2.3 Previous literature

Despite ETFs have emerged quite recently there is a considerable amount of academic research on these types of investment vehicles. Charupat and Miu (2013) did a literature review on modern ETF research. The authors identified three different main strands of research examining different aspects of ETFs, namely ETF pricing efficiency, tracking ability and performance, and finally the effects of ETF on its underlying assets. Of these different main strands, ETF pricing efficiency is discussed in this thesis. Studies regarding ETF pricing efficiency aim to examine how ETFs are priced and how effectively ETF arbitrage mechanism performs. Characteristics of these kind of studies are that they usually involve testing whether ETFs trade close to their NAVs, and in addition, how quickly possible deviation disappears. (Charupat and Miu, 2013)

2.3.1 ETF pricing efficiency in general

Many of ETFs have emerged only recently. However, the interest towards ETF pricing mechanics dates to inception of first ETFs. In their early study on ETF pricing dynamics, Elton et al. (2002) examined SPDR or the so-called "Spider" fund tracking the performance of S&P 500 index. Spider can be considered as a prototype of modern ETF for its purpose and functionality. Elton et al. (2002) tested whether the creation/redemption process can maintain Spider's trading price close to its NAV. They found Spider being rather efficient in terms of pricing. On yearly basis the price deviation from NAV was only 1.8 bps (basis points), which is significantly less than with any of close-end funds to which Spider was compared to. Any deviation in Spider's pricing disappeared within a day when tested with autoregressive framework. The obtained results indicated efficiency in arbitrage mechanism in case of Spider.

Since Spider, also other ETFs emerged tracking more exotic assets. One early category of ETFs were international ETFs tracking equity indices of different countries. Here shares of ETF could be traded in stock exchanges in the US while ETF offered exposure to international assets, for example to Japanese equity indices. Engle and Sarkar (2006) and Delcoure and Zhong (2007) were among the first ones to analyze a sample of these ETFs. Engle and Sarkar (2006) compared pricing efficiency of domestic and international ETFs. Regarding domestic ETFs, the results were consistent with early findings of Elton et al. (2002). ETFs tracking domestic indices were priced efficiently compared to their NAVs. Funds exhibited generally small premiums which disappeared in minutes. As opposite, international ETFs had larger deviations persisting even days. The obtained result regarding domestic and international ETFs suggest that there is more pricing inefficiency incorporated within international ETFs. Similar results were obtained by Delcoure and Zhong (2007), except that they studied only a sample of international ETFs. These results indicated higher inefficiency in the main arbitrage mechanism performed by APs, implying that the process takes longer to correct pricing inefficiencies in case of international ETFs. Charteris (2013) and Petäjistö (2017) further imply, that increased and persistent premiums are also economically significant, as these types of premiums are more interesting from the perspective of arbitrageurs trading in the secondary market.

Consistent with Engle and Sarkar (2006) also more recent research has also found similar results on the general ETF mispricing. For example, Petäjistö (2017) and Hilliard (2014) used a larger sample of US domiciled ETFs with consistent results. Both authors also included other than equity ETFs in their analyses which widened the coverage of different ETF categories in comparison with earlier research. Both Petäjistö (2017) and Hilliard (2014) found that not only international but also bond-based ETFs to have larger premiums compared to domestic equity ETFs. In both studies two former types of ETFs are argued to have more barriers on ETF arbitrage resulting in larger pricing inefficiencies. For bond ETFs, the findings were explained by illiquidity of underlying assets and price quoting, while international ETFs are prone for pricing mismatches resulting from time-zone differences between trading hours of ETFs and their underlying assets.

Petäjistö (2017) further argued that overall ETFs are, on average, priced correctly. To detect mispricing, premium standard deviation or premium volatility should also be examined more closely. Reasoning behind the argument is that despite being correctly priced on average, ETFs can exhibit large premium volatility leading to peaks of mispricing. Certain types of ETF such international ETFs exhibited larger premium volatilities compared to, for example, domestic ETFs. It is notable that results of Petäjistö (2017) are consistent with those of Elton et al. (2002) and Engle and Sarkar (2006) in spite of differences between sample periods, thereby indicating that despite ETF industry is developing, the mispricing is a persistent phenomenon. Moreover, academic research suggests international ETFs to be more mispriced than domestic alternatives. This consensus seems to be supported by all the above-mentioned studies.

2.3.2 ETF pricing efficiency outside the US market

Vast amount of ETF research is conducted with the US domiciled ETFs as the US has the largest ETF markets. Given increasing popularity of ETFs, academia has also started to address ETF efficiency outside the US. Several studies examine ETF efficiency in other markets. For example, Charteris (2013), Purohit and Malhotra (2015) and Tripathi and Garg (2016) have addressed the situation in frontier exchanges with equity ETFs. Results suggest inefficiencies to also occur in smaller exchanges. Charteris (2013) found some South African domiciled ETFs to exhibit significant premiums that persisted more than one trading day indicating some form of pricing inefficiency. Purohit and Malhotra (2015) tested ETF pricing efficiency in India and found premiums of Indian ETFs persisting longer than in any other study. The average persistence was three trading days, and in case of some individual funds, even five days. The results indicate heavy inefficiency and poor functionality of ETF arbitrage mechanism in India.

Tripathi and Garg (2016) compared ETF pricing efficiency between different countries. The selected countries in the analysis were US, Japan, UK, Australia and India. From each country handful of ETFs tracking main stock indices were taken under analysis. The results indicated US to be the most efficient regarding both premium magnitude and premium persistence. By contrast the Indian ETFs exhibited large and persistent premiums indicating inefficiency. Other national ETF markets were deemed to be efficient despite they did not achieve the same level of efficiency as the US markets. The authors discussed potential reason to be lower level of market participation which leads to more ineffective arbitrage process. Regarding the studies of smaller exchanges, it should be noted that they feature only handful of ETFs which mainly track equity indices. As equity ETFs dominate markets especially in smaller exchanges, a comprehensive research including multiple ETF categories may not be possible to conduct in smaller exchanges.

2.3.3 Determinants behind premiums and their magnitude

ETF premiums are important concept as they are used as a measurement to detect mispricing. Most researchers use the calculated premiums to conduct further analysis to examine mispricing. Numerous studies on ETFs aim to explain what causes ETFs to trade at premium relative to fund’s NAV. Current academic research identifies number of reasons causing larger premiums such as staleness of NAV, ETF specific characteristics or overall market sentiment. Common to all these factors is their relation to ETF arbitrage. All three factors can affect how efficiently APs can conduct ETF arbitrage process. When arbitrage is harder for APs to execute, premiums tend to increase.

Many academics estimate ETF premiums simply as deviation of fund’s trading price from its NAV. However, some researchers argue that this simple method is not enough to capture true mispricing due the issue of stale NAV. The concept of stale NAV is particularly important in cases of international ETFs. The issue was first discussed by Engle and Sarkar (2006). They argued that mispricing of international ETFs is at least partly caused by time-zone mismatch of ETF trading hours and reporting of NAV. An example to illustrate stale NAV issue could be as follows: consider a US domiciled ETF tracking the Japanese Nikkei index. Trading with securities included in the Nikkei index ends in Japan before trading in the US begins. Moreover, NAV based on the Nikkei securities is calculated based on closing prices at the end of trading day in Japan. As a result, NAV remains “stale” when ETF shares begin to be traded in the US market. This leaves ETF prices in the US fluctuating and adjusting to new information while NAV does not change as long as the Japanese stock market remains closed. This type of time-zone mismatch may result in seemingly larger premiums in cases of ETFs with underlying assets in foreign markets. The stale NAV issue can also apply for ETFs having illiquid securities underlying them. Moreover, the stale NAV issue can be partly used to explain why research has found international ETFs and some bond ETFs to be more prone to mispricing. (Petäjistö, 2017)

Only few studies have proposed methods for adjusting to stale NAV issue while there is still lack of unifying method. Delcoure and Zhong (2007) made comparisons of the known models to correct staleness issue in their study of 20 international ETFs domiciled in the US. Equation 1 represents a model originally developed by Engle and Sarkar (2006). The model estimates true premiums from regression residuals. Variable x_t is a set of exogenous variables used to explain true premium. Delcoure and Zhong (2007) used S&P 500 returns and currency exchange rate of US and ETF target country in model as exogenous variables.

$$\ln(P_t) - \ln(NAV_t) = \alpha\Delta \ln(NAV_t) + \boldsymbol{\Phi} x_t + u_t \tag{1}$$

Equation 2 is derived from mutual fund industry and extended to the field of ETFs by Delcoure and Zhong (2007). Originally, the model was used in the study of Goetzmann et al. (2001) to overcome similar issues with mutual funds investing to foreign assets. In their model, variable β_i stands for slope coefficient of separate regression estimated with next day NAV return and instrumental variable Z_t, represented by S&P 500 returns in their study. Mathematically the authors present the following model:

$$NAV'_{i,t} = NAV_{i,t}(1 + \beta_i Z_t) \tag{2}$$

Both models thus assume that true NAV can be explained by movements of selected variables during non-synchronous trading hours. Comparison between premiums estimated with straightforward deviation of ETF trading price and NAV and those implied by the above-presented models provides different estimates. However, differences are rather small. Regardless of estimation method, all premiums provided qualitatively similar results when they were used as dependent variables in different analyses. (Delcoure and Zhong, 2007)

Levy and Liebermann (2013) extended the knowledge of stale NAVs by examining intraday patterns of international ETFs trading in the US. Based on the results, the authors argued that there is a significant difference in price discovery between synchronous and non-synchronous trading hours. When ETFs trade at the same time as their underlying assets, price formation is driven by NAV returns. When NAV is not updated during non-synchronous hours, ETF price formation begins to follow S&P 500 returns. These findings verify a rationale behind the above-described models using market indices as benchmarks for stale NAV correction. However, the reported results of Levy and Liebermann (2013) also indicate potential overreaction of international ETFs to the US market returns during non-synchronous trading hours. This, in turn, presents also a potential issue from the perspective of the above-mentioned models. Reflecting the findings of Levy and Liebermann (2013) to Tse and Martinez (2007) yields bit mixed results regarding the matter. Tse and Martinez (2007) found similar indications in their study of 24 international ETFs trading in the US. ETFs had a strong correlation with the US markets. However, the authors reported more conservative results than Levy and Liebermann (2013) as they claimed that despite pricing of international ETFs correlates with the US markets, price discovery and information transmission mainly occurred during trading hours of underlying assets in their local markets. The authors concluded that the examined funds reflected all fundamental information of their underlying assets.

Petäjistö (2017) is the most recent author to offer a possible alternative solution to adjust for stale NAVs. He proposes peer-group method as solution for the issue. The method groups ETFs into sub-groups based on their underlying assets. For example, ETFs tracking S&P 500 are a single group. Petäjistö (2017) then estimates premium as difference of ETF price and peer-group average price. The method discards NAV completely and uses peer comparison for detecting pricing inefficiencies. Moreover, method demands larger sample to form peer-groups, and thus it may not be usable with smaller samples or with ETFs having no peers in the sample. Petäjistö (2017) found that estimating premiums using both unadjusted and adjusted methods provided qualitatively similar results within different ETF categories. International and bond-based ETFs remained the most mispriced ETFs, although magnitude of premiums and premium volatility decreased if estimated with peer-group method.

Despite the attempts to address the stale NAV issue many academics such as Hilliard (2014), Ackert and Tian (2008), Ben-David et al. (2012) and Fulkerson et al. (2014) estimate premiums straightforward as deviations between ETF trading price and NAV. All mentioned studies exhibit a lack of stale NAV correction, despite international ETFs are present in the used samples. A lack of unifying and commonly accepted method may be a reason why some researchers still use simpler estimations when calculating premiums despite staleness in NAVs is found to influence mispricing. Furthermore, the presented correction models have been only applied for ETFs trading in the US markets.

Apart from the stale NAV issue, also other ETF characteristics have been documented to partly explain premiums and their magnitude. Delcoure and Zhong (2007) and Ackert and Tian (2008) are among the first to explain why ETFs trade at premium. Both pair of authors used sample of international ETFs in their panel data-based analyses. Delcoure and Zhong (2007) found that absolute values of premiums have a positive relationship with bid-ask spread, trading volume and currency exchange rate volatility (between the US dollar and the currency of an asset underlying ETF). These can be viewed as fund specific barriers to ETF arbitrage increasing magnitude of mispricing. Higher bid-ask spread, trading volume and currency rate volatility increase uncertainty and costs from AP's perspective. Delcoure and Zhong (2007) also found negative relation of institutional ownership to premiums. As institutional ownership increases arbitrage becomes more effective.

Ackert and Tian (2008) also focused on explaining ETF premium levels of both US and international ETFs. The authors used mainly liquidity, size and momentum-based factors in their analysis. Especially liquidity was under more precise consideration. The authors reported an inverted U-shaped relationship between premiums and market liquidity, as after certain level of liquidity is reached, premiums are decreased. Based on the results, the international ETFs had lower liquidity than the US ETFs. This was argued to cause the higher mispricing of the international ETFs compared to the US ETFs in the study.

Early studies of previous authors have been expanded by more recent attempts to explain ETF premiums and premium magnitude. Fulkerson et al. (2014) tested mainly liquidity and return factors with a sample of US domiciled bond ETFs. The results indicated that liquidity is an important driver of premium level also in pure bond-based ETF sample. Badenhorst (2017) modelled determinants behind premium magnitude in South African ETFs based on annual data. Annual metrics in this case impose limitations as most studies utilize more frequent data. This limitation is also noted by Badenhorst (2017) himself. However, he argued that positive and negative premiums should be considered as separate subjects and modelled separately in contradiction to previous research. He found that factors affecting magnitude of positive premiums may be different than factors affecting negative premiums. For example, investors are willing to pay premium for ETFs investing in illiquid assets, but similar linkage does not naturally exist with negative premiums.

One of the most recent studies on ETF premiums is Picotti (2018), who aimed to explain relation between ETF premiums and ETF liquidity segmentation. His findings suggest that investors are willing to pay premium for ETFs investing in underlying assets that are illiquid. Reasoning is that ETFs offer easy access for such asset classes that could be otherwise limited for many investors. Conclusion is similar than offered by Badenhorst (2017) and Hilliard (2014). This reasoning of willingly paid premiums offers possible explanation why international and bond-based ETFs are more mispriced even when the stale NAV issue is controlled. Another important finding of Picotti (2018) is that the age of ETF is related to premium level. Younger ETFs tend to have higher positive premiums. Picotti (2018) argues that this derives from the higher liquidity benefits associated with newly issued ETFs.

Apart from ETF characteristics driving premiums, overall market conditions have documented to also have an influence. Regarding market level determinants, there are only handful of studies using them as explanatory variables in econometric models. Petäjistö (2017) and Ben-David et al. (2012) are among the few to explain ETF mispricing with overall market sentiment. Findings on overall market variables suggest that when uncertainty increases ETFs tend to be more mispriced. Both Ben-David et al. (2012) and Petäjistö (2017) support this point of view. CBOE VIX index used in both studies was found to have a positive relationship with mispricing of ETFs traded in the US. Also, capital constraints measured with TED spread (difference between 3-month Libor and US 3-month T-bill) were found to have positive effect on ETF mispricing in both studies. The rationale lies again in arbitrage process: when markets exhibit higher volatility and uncertainty, there is less arbitrage capital willing to attempt ETF arbitrage. This, in turn, allows ETF prices to further move away from NAVs. Similar reasoning is offered by Kreis and Licht. (2018). Despite they focused on trading against mispricing, they found that the spread between the highest premium and lowest the discount is magnified during times of financial crises. Also, Delcoure and Zhong (2007) documented relationships between premium magnitude and crises. For example, the Asian (1997) and Russian (1998) financial crises and the 9/11 tragedy in the US support relations of higher premiums and extreme market conditions.

2.3.4 Trading strategies against mispricing

EMH by Fama (1970) states that investors should not be able to use public information to make excess returns and "beat the market". In this light an individual daily ETF premium is not necessarily a violation against EMH. Moreover, if EMH holds, using this type of information of ETF premiums should not yield any excess returns. Consequently, many researchers test ETF mispricing with active trading strategies to see whether abnormal returns can be obtained. One type of these strategies is a long-short trading against mispricing.

One of the first studies to incorporate trading against mispriced ETFs was Jares and Lavin (2004). Their study involved testing inefficiencies with the two US domiciled ETFs tracking performance of stock indices in Japan and Hong Kong. They found that ETF premiums exhibit mean reversion. Mean reversion suggests that when ETF price drifts away from funds NAV it is usually followed by movement towards NAV soon after. Similar mean reversion process predicting future returns was later documented in several studies: for example, in Rompotis (2010), Kreis et al. (2016) and Charteris (2013). The mean reversion process has potential to be exploited in trading due to its predictive abilities. Jares and Lavin (2004) formed simple trading strategy around this finding:

Premium rule: if ETF price is more than NAV – take a short position on ETF shares.

Discount rule: if ETF price is less than NAV – take a long position on ETF shares.

This type of long-short strategy is based on the fact that if trading at premium, ETF price is expected to revert towards its NAV. The reverse holds for ETF trading at discount or negative premium. Investor taking appropriate positions according to ETF premiums can capture this price movement and make excess profit over buy-and-hold strategy. Cumulative profits of Jares and Lavin (2004) provided evidence that their strategy indeed could beat the buy-and-hold benchmark in cases of both ETFs. Active trading on the Japan-based ETF generated 542% while the Hong Kong-based ETF generated 12119% of total cumulative returns. Trading days reported were 1361 days for the Japan-based ETF and 1364 days for the Hong Kong-based ETF between years 1996 and 2001. In their study Jares and Lavin (2004) report cumulative return for each year individually in addition to the previously-mentioned total cumulative returns. Assuming 252 trading days in a year, annualized total returns for the whole sample period based on the reported total cumulative returns are approximately 41% for the Japan-based ETF and 143% for the Hong Kong-based ETF. The results are quite extreme as buy-and-hold benchmarks could produce only negative cumulative returns during the same period. For the Japan-based ETF cost of 0.52% per transaction must be introduced before return of long-short trading matches the buy-and-hold benchmark. For the Hong Kong-based ETF the same estimate is 1.082% per transaction. However, applying transaction costs naturally decreases excess returns of the long-short strategies over the buy-and-hold benchmarks. A filter of choice can be applied to activate trades only when premiums are above the selected filter. Jares and Lavin (2004) applied +/- 0.10$ filter for their strategy. The authors found this type of strategy to increase returns while including transaction costs as filter cuts down number of trades. The results violate EMH even in its weakest form, as trading is based solely on information of historical prices.

Since Jares and Lavin (2004), other researchers have also adopted similar long-short trading to test whether mispricing of ETFs can be exploited in a profitable manner. The profitability of active trading has been tested with several types of ETFs. More recent research based on exploiting ETF mispricing in trading involves, for example, Fulkerson et al. (2014), Petäjistö (2017), Kreis et al. (2016), Kreis and Licht (2018) and Charteris (2013). Common conclusion to all these studies is potential to exploit mispricing with active trading.

As Charteris (2013) extended ETF literature to South African markets she also tested whether a similar trading strategy to that employed by Jares and Lavin (2004) could produce profits. She reported that three out of seven funds under analysis were able to produce excess profits even after taking account of transaction costs. As the filter for both long and short trading signals, Charteris (2013) calculated absolute value of long-term premium.

Fulkerson et al. (2014) analyzed pricing efficiency with a sample featuring 140 different bond-based ETFs domiciled in the US. Regarding their long-short trading strategy Fulkerson et al. (2014) found that bond ETFs were able to produce excess returns over 11% on yearly basis. Their approach included dividing ETFs into deciles based on ratios of ETF price to NAV and taking appropriate positions for holding over next month. The documented results by using monthly data indicated that trading against mispricing can be profitable also on monthly level, rather than purely daily level. The authors validated the achieved returns with a bond factor model. Regressing returns against factors controlling for bond markets produced significant positive alpha.

Kreis et al. (2016) did their research with a sample containing handful of US domiciled ETFs on Latin American equity indices. Their findings are also consistent with other studies. Long-short trading provided excess returns in their sample before transaction costs. Instead of devising signals for trading positions, Kreis et al. (2016) simply took only two daily positions: long on ETF having the largest negative premium and short position for ETF having the largest positive premium. The CAPM and the three-factor model of Fama and French (1993) were used to test the significance of excess returns. The daily excess return was reported to be 0.5% before transaction costs.

Kreis and Licht (2018) applied similar methods for ETFs trading in Europe. They studied 19 ETFs trading at XETRA and following different sectors of Eurostoxx index. Findings for the European ETFs indicate that gross-excess profit is possible to be acquired in XETRA. Moreover, the authors report that if transaction costs are added the net excess return is positive only in years 2008-2010. Kreis and Licht (2018) argue this to be due to the higher volatility during the financial crisis. When market uncertainty is increased, premiums tend to be larger for arbitrageurs to capture. When market volatility normalizes after a crisis period, the premiums decrease while profit potential in mispricing strategies decreases in conjunction. According to Kreis and Licht (2018), premiums were too low to enable profitable trading after transaction costs. These results qualitatively validate results regarding effects of overall market sentiment on ETF pricing efficiency.

Petäjistö (2017) also tested long-short trading as part a of his research. An extensive sample of 586 US domiciled ETFs also produced excess returns. Based on a large sample Petäjistö (2017) was able to document whether there are differences between ETF categories. The results showed that excess return is increased if the strategy is applied for ETFs more prone to mispricing. When returns were regressed within the CAPM framework, significant positive alpha was also found. Similar results were obtained with the Fama-French three-factor and the Carhart (1997) four-factor models. The largest annualized alpha was over 16% when strategy was applied solely to the most mispriced ETF categories such as international ETFs and some bond ETFs. These results did not take account transaction costs. Petäjistö (2017) also discussed that strategy can be executed in profitable manner even with costs included. However, detailed description or results of this analysis were not included in the paper.

There are also other types of long-short strategies tested. For example, Davidson (2013) formed a strategy where he would go long on ETFs and short underlying assets based on signals given by different market factors. Selected factors were CBOE VIX, S&P 500 VIX Short-term futures, S&P 500 TR index and Russell-Axioma index as liquidity proxy. Using a sample of US domiciled ETFs following domestic securities he found annualized Carhart alpha of 7.22%. Rather similar approach was used by Simon and Sternberg (2005). They studied three ETFs based on German, UK and French securities and found profitability on going long or short either outright or taking offsetting position using Spider (SPY) as a hedge. For signaling they used premiums in accordance with previously-described studies. This type of approach yielded profits from 0.25 to 0.50% per day before transaction costs.

Based on reviewed research, long-short strategies indicate important notations in conjunction of the results provided by other previously-mentioned strands of ETF research. First, it is evident that ETF categories more prone to mispricing exhibit higher and more persistent premiums that can be exploited better in terms of trading against mispricing. This stems from the fact that larger premiums have stronger mean reversion effects while higher persistence provides time to take appropriate trading positions from the investor's perspective. Second, even different types of strategies based on long-short framework seem to be able to generate excess profits. Notably different authors use different filters, portfolio balancing and ETF samples in their trading. Despite differences profitability remains. Profits also seem to be magnified during the times of market distress. Notably most of the samples include financial crises of 2008 in the analysis.

The mean reversion process of ETFs is qualitatively a challenge against EMH and even its weakest form since relatively simple trading rules based on historical data often lead to excess returns. Marshall et al. (2013) document arbitrage opportunities for two S&P 500 ETFs. Usage of intraday data suggests that even the most liquid ETFs can exhibit profitable arbitrage opportunities intra-daily. These opportunities might not be available at daily level, as most of the other reviewed studies suggest. Moreover, intraday trading presents another interesting strand for exploiting mispricing with active trading.

3. DATA AND METHODOLOGY

3.1 Data

Data for ETFs analyzed in this thesis is obtained from Thomson Reuters Datastream. Searching for ETFs listed in LSE results in over 2000 potential funds. However, some filtering is needed. Many ETFs trade in multiple exchanges and even have multiple quotes in LSE at different currencies. Filters are set to include only pound denominated ETFs domiciled in the UK or Ireland and currently having primary quote available in LSE. These filters should remove any duplicates of ETFs listed in multiple currencies and funds that are domiciled in foreign countries. Filters should therefore ensure that focus is on ETFs that are available for regular traders in LSE, but without reducing overall sample size smaller than used in the reviewed literature.

After the previously-described filtering, a sample of 301 ETFs is left for a time period from 2.1.2014 to 29.12.2017. Observations on holidays and weekends or other days when trading is unavailable are filtered out leaving total number of days included at 1011. Only ETFs having data for over 252 days were included in the sample. Variables obtained for each ETF includes closing price, NAV, closing bid and ask quotes, market value in pounds and inception date. However, not all 301 ETFs have complete data available for all the variables. The sample was further narrowed and ETFs having less data than 95% of possible maximum during the period were not included. In addition data for daily trading volume was obtained. However, for many ETFs this data contained lots of missing values and thus it is not included in the 95% filter. Final sample consists of 201 ETFs investing in different types of securities. To include cross-sectional differences in analysis, ETF investment styles were obtained from Morningstar. The acquired sample can be seen consistent with the reviewed ETF literature in terms of number of ETFs and length of sample period. Market variables that represent overall market conditions in the UK including FTSE All Shares return index, FTSE 100 VIX, 3-month Libor rate and 3-month UK treasury-bill rate were obtained from Datastream. In addition, proxies for market portfolios were obtained from data library of Kenneth French.

Table 1 presents descriptive statistics for the sample and for all main variables to be used in further analyses. Variables are divided into groups based on models they are used in. The statistics reveal that none of the variables is normally distributed. Most variables exhibit excess kurtosis (over 3 of normal distribution). This type of kurtosis leaves most variables to be leptokurtic. Regarding skewness most variables have positive values. Therefore, tests for normality using skewness and kurtosis should reject normality in case of each variable. This is confirmed with Jarque-Bera test. From statistical perspective normal distribution is assumption behind statistical models. Rejecting normality in case of all variables leads to violation of this assumption in the models.

Table 1: Descriptive statistics. Table present descriptive statistics for each variable used in the following empirical models.

Variable	Mean	Std.	Min.	Max.	Skew	Kurt.	Obs.
Price (GBX)	3942.14	4173.61	215.55	39365	2.16	8.83	171354
NAV (GBX)	3946.23	4173.58	215.45	32340.18	2.16	8.81	170923
Premium (%)	0.05	0.48	-8.20	17.54	1.64	67.56	170923
ETF level variables							
AbsPremium (%)	0.29	0.38	0	17.54	6.92	136.83	170923
Bid-Ask spread (%)	0.37	0.46	0	17.92	5.86	68.38	171263
NAV return volatility (%/5-day rolling std.)	0.81	0.58	0	11.45	2.73	20.01	169914
Market value (million GBP)	451.52	995.21	0.03	10033.15	4.52	28.12	171269
Logarithm of Market value	4.48	2.05	-3.51	9.21	-0.22	2.55	171269
Age (Years)	6.05	3.76	1.00	18.00	0.54	2.64	171356
Market level variables							
Cross-sectional premium dispersion (%)	0.41	0.20	0.16	2.06	2.75	15.27	1011
FTSE 100 VIX quote	14.70	4.40	6.19	32.48	1.40	5.03	1011
FTSE All shares return (%)	0.03	0.83	-4.53	3.47	-0.21	6.11	1011
UK TED spread (%)	0.17	0.06	0.06	0.38	1.06	3.70	1011

3.2 Methodology

The reviewed literature suggests econometric methods which have been previously applied in the field of ETF mispricing. Studies rarely focus on testing ETF premiums with just one method but rather use variety of different tests to provide a comprehensive analysis. Similar approach is adopted here. ETFs are analyzed with different types of models to answer different research questions. Table 2 summarizes the formed null hypotheses, as well as statistical models used for hypothesis testing. Robustness analyses are done by using stale NAV adjusted premiums instead of unadjusted ones and by dividing the sample into sub-periods. In the sub-period analysis unadjusted premiums are used.

Table 2: Hypotheses. Table provides the formed null hypotheses. Also, model framework, related reviewed literature and the related research questions in this thesis are presented.

Hypothesis	Model	Related literature	Related research question
H1: ETFs do not trade at significant premiums	One-sample t-test	Charteris (2013)	1. & 4.
H2: ETF does not trade at persistent premiums	AR with 1 to 5 lags	Charteris (2013) Rompotis (2010)	1. & 4.
H3: ETF level factor has no relation to premium magnitude.	Panel regression	Delcoure and Zhong (2007) Fulkerson et al. (2014)	2. & 4.
H4: Market factor has no relation to premium dispersion.	OLS regression	Ben-David et al. (2012) Petäjistö (2017)	2. & 4.
H5: ETFs show no evidence of inefficiencies to be exploited in trading.	Long/short trading CAPM	Charteris (2013) Kreis and Licht (2018) Petäjistö (2017)	3. & 4.

3.2.1 Modelling ETF premiums

ETF premium is a key variable in empirical part of this thesis. Here premiums are modelled with equation 3. Premium is simply a deviation between ETF trading price and NAV of the given day. This type of model is the most common way to estimate premiums even though it is not taking staleness of NAV into consideration.

$$PREM_t = \frac{PRICE_t}{NAV_t} - 1 \tag{3}$$

Stale NAV correction models could be other and more proper solution to estimate premiums. However, literature does not offer evidence of ETF premiums to be dependent on the UK or other European markets during non-synchronous trading hours similar than in the US. As there is a lack of empirical evidence in adjustments in Europe, the stale NAV adjusted premiums are used as one form of complementary analysis for the results based on premiums of equation 3. Furthermore, it is possible to conduct preliminary analysis whether this type of approach is usable with ETFs traded in LSE as mismatch from time-zones is different than for traders in the US. From UK investors' perspective Asian market do close before trading begins in the UK. This is similar mismatch as encountered in research on the US domiciled ETFs. Main difference in European setting is that UK has several hours of simultaneous trading with the US markets before the UK markets close. In this situation it would be suitable to use intraday NAVs (iNAVs) released by listing exchange for the US-based funds (Kreis and Licht 2018). Since there are overlapping trading hours, iNAVs calculated around the UK close for ETFs investing in the US assets should take account of the staleness issues most accurately. Unfortunately, for ETFs trading in LSE, iNAVs are not available and therefore correction method could be applied instead.

Correction for staleness is done based on the model of Engle and Sarkar (2006) with specification of Delcoure and Zhong (2007). Later these premiums are referred to Engle-adjusted premiums. Since iNAVs are unavailable, similar correction is applied for the Asian and US based funds alike. As a robustness check, the same method is used also for the

European ETFs. Premiums and premium volatilities of these ETFs should not change much since there is none or very little timing mismatches.

Equation 1 presented in theoretical framework is used for premium specification and FTSE All shares return index is used as $\boldsymbol{\Phi} x_t$. Thus, corrected premiums are determined by change in NAV and change in the benchmark index. This follows approach of Picotti (2018) who used the US market portfolio for similar adjustment when studying the US domiciled ETFs. The same method is also used for different ETFs, for example, bond and equity ETFs alike as Picotti (2018) did. The method should decrease volatility of adjusted premiums in cases of ETFs suffering from staleness of NAV as comparison to unadjusted ones.

Appendix 1 provides comparison between regular premiums and the Engle-adjusted premiums. Based on comparison between unadjusted and adjusted premiums the effect of the adjustment is similar to that documented by Delcoure and Zhong (2007). Both cross-sectional average premium and standard deviation of premiums do decrease after adjusting premiums with the employed model.

Large premiums may be a result of measurement errors in data. Following Petäjistö (2017) and Picotti (2018), every premium over 20% is inspected closely. The results of inspection are similar with the ones in the last-mentioned papers, and every premium over 20% are filtered from the sample as they are errors in data. Also 10% filter was examined, but premiums on that level are legitimate and thus included in analysis. Also, a handful of clear data errors regarding premiums were replaced by information provided by official releases of LSE. In addition, other variables were inspected for possible clear errors in data and such error observations were removed from sample.

3.2.2 Modelling mispricing and its persistence

Discovering mispriced ETFs relative to fund's NAV can be done on the basis of premiums from equation 3. Average premium can be estimated for each ETF and for each ETF category. One sample t-test can then be used to test whether a premium significantly differs from 0. If t-statistic is significant it indicates that ETFs do trade at positive or negative premium leading to rejection of H1. Petäjistö (2017) suggests that mispricing may also be a result of premium volatility. ETFs can be trading close to their NAVs on average, but premium volatility may cause a periodical mispricing. Following this intuition, premium volatility with 95% confidence band is also calculated for each ETF category.

Testing premium persistence is important. If premiums do not persist, investors cannot capitalize mispricing. Persistence analysis can be done with AR or autoregressive model within ARMA-framework. The purpose of AR-models is to estimate dependent variable with its own lagged values (Brooks 2014, 259-260). Regarding ETF premiums AR-model can be used to estimate whether premiums of previous days would have explanatory power over today's premium. Following equation 4 describes the model used.

$$PREM_t = \alpha + \beta_1 PREM_{t-1} + \beta_2 PREM_{t-2} + \beta_3 PREM_{t-3} + \beta_4 PREM_{t-4} + \beta_5 PREM_{t-5} + \varepsilon_t \tag{4}$$

In equation 4 variable $PREM_t$ represents ETF premium in time t. The right-hand side variables, except the residual term, are lagged values of these premiums. Here interest lies in beta coefficients. If betas are statistically insignificant, premiums disappear within one trading day. If only the first beta is significant it means that premium persists for only one day. First two significant betas would imply at least two-day persistence. Here H2 is rejected if at least one beta is significant. Charteris (2013) argues that modelling should start from inclusion of just one lag and then continued with additional lags until the last beta becomes insignificant. Following this suggestion, modelling is stopped in lag two if first beta is significant but the second is not. Modelling is also stopped if previous betas turn insignificant after introducing additional lags. Each ETF is modelled individually using five lags as maximum. Also, modelling is extended to panel data setting, where ETFs in each category are modelled together with five lags. This is done as a robustness check to find out whether there are differences between analyzing ETFs individually or as a larger sample.

3.2.3 Determinants for mispricing magnitude with panel data and OLS

Following approaches of previous studies, determinants driving ETF mispricing are explored with linear models. Here both traditional OLS and panel data regressions are used. ETF data obtained has both cross-sectional and time dimensions allowing possibility to do richer analysis. For example, standard linear regression estimated with OLS (ordinary least squares) can only capture either cross-sectional or time-series variations in the model. With panel data controlling unobserved heterogeneity of cross-sectional variables over time is possible. Given the benefits of panel data regression, it is a reasonable framework to study a larger sample of ETFs. (Brooks 2014, 526-529).

The overall setting of panel data can be presented in equation 5. Here dependent variable y_{it} is explained with explanatory variables x_{it}. As combination of time and cross-sections panel data approach allows increased degrees of freedom. Sample size in equation 5 is number of cross-sectional variables multiplied by number of time-series observations given N×T. (Brooks 2014, 526-527; Baltagi 2005, 11)

$$y_{it} = \alpha + \beta X_{it} + u_{it} \tag{5}$$

Here α is constant, β is slope vector of explanatory variable, X_{it} is observation of cross-sectional variable (i) at given time (t) and u_{it} is error term. Error term u_{it} can be further divided according to equation 6.

$$u_{it} = \mu_i + v_{it} \tag{6}$$

Here μ_i is individual specific effects and v_{it} is remainder disturbance. μ_i can be though as term for capturing cross-sectional effects on dependent variable that do not vary over time. v_{it} can be described as usual disturbance term used in regression models. (Brooks 2014; 526-529)

There are mainly two types of panel data models commonly used in financial econometrics: fixed and random effects models. Despite both models operating in panel data setting, there are differences in approaches. Fixed effects model is used, if there is individual specific effect that is not included in regression and is correlated with X_{it}. Fixed effects model assumes μ_i to be fixed while v_{it} is independent and normally distributed for all cross-sectional variables and time periods. In this model differences between cross-sections are captured with intercept. Notably also time-fixed effects can be introduced instead of cross-sectional fixed effects if assumed that effects vary across time while being constant cross-sectionally. Combining both types of effects together in same model is also possible and is known as two-way error component model. Moreover, Baltagi (2005) suggests that fixed effect model is appropriate approach when focus is on specific set of N and our inference is restricted to behavior of selected N. (Baltagi 2005, 12-19; Brooks 2014, 528-532)

Random effect model assumes unobserved individual heterogeneity to be uncorrelated with included variables. In this setting X_{it} is independent from μ_i and v_{it} for all i and t. μ_i is used to capture unobserved heterogeneity as intercept is the same for all cross-sections. Compared to the fixed effect, random effects model is more appropriate if N is drawn randomly from large population. (Baltagi 2005, 12-19; Brooks 2014, 536-537)

To choose a proper estimation method statistical tests need to be applied. The Hausman test is first used to derive whether fixed or random effects model is the proper one. H0 of the Hausman test can be though as a test of whether fixed and random effects coefficients are the same. If H0 is rejected fixed effects is suggested. If H0 is not rejected, random effects model should be considered. However, further testing is needed. Fixed effects model should have a same constant for all units. This H0 can be tested with F-test for no fixed effects. Rejection points further towards fixed effects when no rejection suggests using pooled OLS method. Similarly, random effects need further testing. The Breusch-Pagan test can be used to test whether there are no random effects. If this H0 is rejected, random effects are suggested. If H0 indicates no presence of random effects we can use pooled OLS. To find proper method for panel data both fixed and random methods are used and selection of more appropriate model is derived from the described tests. (Baltagi 2005, 12-19; 59)

Following models are estimated in this thesis. Equation 7 is a model where premium magnitude is explained with ETF characteristics. The model follows the previously-described panel data approach. Equation 8 uses market factors as independent variables. This equation is estimated as standard OLS regression because dependent variable premium dispersion does not feature similar cross-sectional element, as it is computed as cross-sectional standard deviation of ETF premiums at given day, following the approach of Petäjistö (2017).

$$|PREM|_{it} = \alpha + \beta_{SPDR}SPRD_{it} + \beta_{VOLNAV}\,VOLNAV_{it} + \beta_{SIZE}SIZE_{it} + \beta_{AGE}AGE_{it} + \varepsilon_{it} \tag{7}$$

$$PREMSTD_t = \alpha + \beta_{VIX}VIX_t + \beta_{MARKET}MARKET_t + \beta_{TED}TED_t + \varepsilon_t \tag{8}$$

Description of variables included in the regression analyses is given in Table 3. Expected sign indicates relationship that variable is expected to have with the dependent variable. Related research implies to reviewed study from which the motivation to use such variable is derived. The last column exhibits more detailed description of the variable behind the abbreviation.

$\boldsymbol{|PREM|_{it}}$: Absolute value of estimated premiums can be used as a measure of premium magnitude. Usage of absolute values does not consider whether premium is positive or negative. For this reason it serves as proper proxy for premium magnitude on either side.

$\boldsymbol{SPRD_{it}}$: ETF arbitrage does include costs. Bid-ask spread measures how costly it is to trade or arbitrage on ETF. As the spread widens APs have higher costs to exploit mispricing. This reasoning indicates positive relationship with premium magnitude as higher costs discourage arbitrage activity.

Table 3: Variables. Table describes the used variables in equations 7 and 8, as well as their expected sign in the empirical models. Also, table presents research behind the motivation to use such variable.

Variable	Expected sign	Related research	Description
Equation 7			
$\lvert PREM\rvert_{it}$	Dependent	Delcoure and Zhong (2007)	Absolute value of premiums from eq. 3.
$SPRD_{it}$	Positive	Delcoure and Zhong (2007)	Bid-Ask spread
$VOLNAV_{it}$	Positive	Ben-David et al. (2012), Hilliard (2014) + own elaboration	Rolling 5-day standard deviation of NAV returns.
$SIZE_{it}$	Negative	Badenhorts (2017)	Log of daily market value (in millions).
AGE_{it}	Negative	Picotti (2018)	Dummy for age over 5y.
Equation 8			
$PREMSTD_{t}$	Dependent	Petäjistö (2017)	Cross-sectional premium dispersion (standard deviation)
VIX_{t}	Positive	Petäjistö (2017)	FTSE 100 VIX quote
$MARKET_{t}$	Negative	Ben-David et al. (2012)	Daily return of FTSE All shares index
TED_{t}	Positive	Petäjistö (2017)	3m Libor – 3m UK T-bill

$\boldsymbol{VOLNAV_{it}}$**:** ETF's fundamental value is linked to the value of its underlying assets. From this perspective it can be more difficult to conduct arbitrage if underlying asset is more volatile. Since ETF arbitrage may require several matching trades on both ETF and underlying assets, increased volatility increases risks in this process. Hilliard (2014) also argues that overall volatility of emerging markets may hinder ETF arbitrage causing these types of ETFs to be more mispriced. The usage of five-day lagged average standard deviation is derived from Ben-David et al. (2012) who found linkage between NAV and premiums to exist on daily level. However, current academic studies also indicate that relationship can also be inverted with the respect that ETF ownership can influence volatility of their underlying assets.

$\boldsymbol{SIZE_{it}}$: Badenhorst (2017) found significance on size explaining premium magnitude on yearly level. However, value of coefficient was almost zero and significance was lost when premiums and discounts were modelled separately. Following these results, possible effects of ETF size to absolute premium can also be tested in LSE to determine whether there is a relationship on daily level. Size can be argued to provide higher liquidity, which should again lead to more efficient arbitrage process. Size is also used in studies of Fulkerson et al. (2014) and Picotti (2018) when modelling level of premiums.

$\boldsymbol{AGE_{it}}$: The age of ETF is argued to have negative relationship with premium level. Picotti (2018) is among the first to consider this type of variable in explaining premiums although he did it from liquidity segmentation perspective. Younger ETFs are argued to exhibit larger premiums as they offer new diversification benefits for investors. Picotti (2018) found ETFs over 5 years old to be trading only a fraction of a premium documented for younger ETFs. By extending his analysis, it can be tested whether similar effect exists with premium magnitude. Thus, the age over 5 years is used as a categorical dummy variable.

$\boldsymbol{PREMSTD_{t}}$**:** Denotes cross-sectional standard deviation of ETF premiums**.** Larger value indicates wider spread between positive and negative premiums leading to larger overall mispricing. Petäjistö (2017) used this measure to test whether overall market conditions affect mispricing magnitude in the US.

$\boldsymbol{VIX_{t}}$**:** ETF premiums are found to be linked to overall market conditions. In the US, increase in CBOE S&P 500 VIX serves as a proxy for increased market riskiness. In UK FTSE 100 VIX can be viewed as a similar measure. Turbulent markets introduce discouragement to conduct ETF arbitrage as process is riskier for APs causing increase in general mispricing.

$\boldsymbol{MARKET_{t}}$**:** Ben-David et al. (2012) argue that also overall market index can serve as a proxy for capital constraints. Downturn in market return causes APs to be more passive regarding ETF arbitrage. In the US, large market indices (S&P 500 for example) are used. In the UK setting, the closest match would be FTSE All Shares return index.

$\boldsymbol{TED_t}$**:** TED spread is a measure of availability of arbitrage capital. When spread increases APs find increased cost for arbitrage capital. Increasing costs also in form of widening TED spread are discouragement to ETF arbitrage activity. Traditionally, TED spread is calculated as difference of 3-month LIBOR and 3-month US T-Bill. In the UK setting, similar TED spread is measured as difference between 3-month LIBOR and 3-month UK T-bill.

Hypothesis testing of H3 and H4 are applied to independent variables as models provide t-statistics for each individual variable. If a variable has a significant effect based on its t-statistics, H3 or H4 depending on the model will be rejected. Based on the rejection, it is possible to address which individual variables contribute to mispricing.

Before estimating models, some assumptions should be considered. Pairwise correlations are estimated to detect possible multicollinearity between independent variables. No multicollinearity is one of the assumptions behind the OLS regression. If such collinearity existed, one of the collinear variables should be left out from the model. Pairwise correlations of the employed variables are presented in Appendix 2. However, the estimated correlations between independent variables are 0.5 at maximum indicating that they can be used in the same model.

Stationarity is another assumption behind time-series modelling. To ensure stationarity, the adjusted Dickey-Fuller test for unit root is applied for each time-series of premiums and absolute premiums. The same method is also applied for the time-series of premium dispersion which is the dependent variable in the equation 11. Akaike information criteria is used to determine appropriate lag length. Of 201 ETFs, a few failed to reject null hypothesis of unit root. Because issues with unit roots were not severe considering the overall sample, these ETFs were also included in the reported results. Qualitatively inclusion or omission of these ETFs does not affect results. Regarding premium dispersion, null hypothesis of unit root was rejected. For panel data models, the Fisher-type panel data unit root test was applied for dependent variables. All performed tests rejected H0 of all panels containing unit roots.

3.2.4 Long-short trading strategy

The third research question covers possibility to utilize long-short trading strategies to exploit daily premiums. Founding rule in this type of trading is as follows:

1. Take long position on ETFs trading at negative premium (discount) relative to NAV.
2. Take short position on ETFs trading at premium relative to NAV.

For returns of each position equation 9 below is used:

$$LONGR_{it} = (P_1 - P_0) / P_0$$
$$SHORTR_{it} = (P_0 - P_1) / P_1 \tag{9}$$

Here $LONGR_{it}$ is return of long positions and $SHORTR_{it}$ is return of short positions. P is closing price for ETF at given day. The return on the strategy is calculated based on initial investment of 1000£. Every day positions are taken into ETFs having premiums over or under selected signal filter. Arbitrary filters have been employed in earlier literature, for example, in Jares and Lavin (2004). Arbitrary filters are tested also in this thesis. For example, considering a filter of +/-150 bps: ETF having premium over the filter of +150 bps would be shorted and ETF having premium below -150 bps would be taken under long position. As universe of potential ETF is broader than in most previous studies, different filters can be employed to focus on most mispriced ETFs and to avoid excess trading. Available capital is divided equally between ETFs having a trading signal at given day. Positions are held one trading day and then strategy starts on following day by taking new positions based on premiums over the selected filter. Thus, the daily return on the strategy is equal weighted average of returns of the taken positions. Profits made on the trades are re-invested. Capital is held at zero-interest account if trading signals are not received at the given day. Henceforth, the strategy based on the employed method is referred as Filter-strategy.

Reviewed studies implicitly assume that investors can take desired positions with same closing prices that are used for signals (for example Charteris 2013, Jares and Lavin 2004, Kreis and Licht 2018). This assumption may bias results as closing prices and official NAVs used in this thesis are calculated and available only after trading ends. In European time-zones the assumption introduces also a look-ahead bias for ETFs having their underlying assets in time-zones where trading continues after LSE closes. As official NAVs are based on the latest closing prices they contain information that investor trading in LSE would not have been able to utilize at the closing of LSE. This applies to, for example, US-based ETFs. In such case investors could take appropriate positions using closing prices only at the beginning of following day when market opens again. If significant new information arrives after trading hours of LSE it is unlikely that investor could take positions based on the previous closing prices. Also, sufficient trading volume in case of some ETFs can add some additional real-life complications and decrease attractiveness of implementing long-short strategies on larger capital (Petäjistö, 2017).

To address these previously-described issues some additional restrictions are used. The Filter-strategy is employed also to different sub-samples including only ETFs having their underlying assets in the markets that close before or simultaneously with LSE. This covers, for example, Asian-based equity ETFs despite time-zone differences. Regarding European-based ETFs, only Frankfurt Stock Exchange is open after the close of LSE. Given that shares traded in Frankfurt feature only a small portion of total assets underlying the European ETFs, these can be considered rather free from potential biases. In caveat, European ETFs included in the sample invest mainly to the whole Eurozone instead of single countries. Theoretically investors should be able to approximate NAVs of both European and Asian ETFs around the close of LSE. In reality, using same day closing prices for trading are also more proper proxy than 1-day lagged closing prices since mispricing can be assumed to disappear quickly. Investors having access to intraday data may also use intraday iNAVs and prices to calculate signals for positions to be taken before market closes. The usage of iNAVs, if available, should thus provide more realistic approach regardless of ETFs used in the analysis.

Also, trading volume might be a concern regarding active trading. The obtained data for daily trading volume contains lots of missing values for many ETFs and it could be an indication that ETF has not been traded during the day. To address this issue, positions can be taken only with ETFs having at least some trading volume available for a day when position is taken and for a day when position is liquidated. This should ensure that trading is possible, despite LSE (2018) reports closing prices based on midpoint of the best available bid and ask quotes even if actual trades have not occurred. For additional pre-caution ETFs need to have daily market value of at least 5 million pounds before being available in the Filter-strategy. This should limit possible effects of very small or newly issued ETFs. Notably Petäjistö (2017) used similar liquidity cutoffs, as his cutoff filters were 10 million dollars of market value and daily trading volume of 0.1 million dollars. As the sample in this thesis is significantly smaller and due to the missing values regarding trading volume, the liquidity cutoffs for the trading are kept more lenient compared to Petäjistö (2017).

Performance comparison between active trading and passive buy-and-hold strategies are based on the Sharpe (1966) ratios and Jensen's (1968) alpha. Buy-and-hold strategy involves equal initial investment to each ETF in the sample. ETFs incepted during the sample period receive initial investment only at the inception point. For other funds, the initial investment is made at the beginning of the sample period. Otherwise, buy-and-hold portfolios are not rebalanced during the holding period. To properly account for a possibility of negative excess returns, the modified Sharpe ratio introduced by Israelsen (2005) is used as a risk-adjusted metric for comparison of trading strategies as follows:

$$SHARPE_i = (Rp_i - Rf_i)/\sigma_i^{(ER\,/\,|ER|)} \qquad (10)$$

where Rp_i is raw return on the strategy, Rf_i is risk free rate, σ_i is standard deviation of excess returns and ER is excess return of the strategy. Both returns and standard deviations are annualized for sake of consistency. Higher value of the Sharpe ratio implies better risk adjusted returns.

CAPM by Sharpe (1964) provides another framework to measure excess returns relative to risk. In the CAPM framework, effects of overall market developments are controlled for. In ETF studies Petäjistö (2017) and Kreis and Licht (2018) have used the CAPM framework to measure abnormal return of trading strategies. Equation 11 presents formulation of CAPM:

$$Rp_t - Rf_t = \alpha + \beta(Rm_t - Rf_t) + \varepsilon_t \tag{11}$$

where variable Rp_t is return of the trading strategy, Rf_t is risk free rate and Rm_t is market portfolio. The resulting excess return is regressed against excess return of the market portfolio. As ETFs in the sample have different types of underlying asset (different equities, bonds and commodities), selection of proper market portfolio is difficult as purely equity-based indices might not be valid options. In this thesis buy-and-hold strategies are employed as market portfolio Rm_t. However, also market portfolios obtained from data library of Kenneth French are employed as a robustness analysis. The UK 3-month T-bill is used as risk free rate. If a long-short trading strategy beats buy-and-hold benchmark and generates positive alpha in the CAPM framework H5 is rejected. Such an outcome is a violation against EMH and particularly against its weakest form as the trading strategy is based purely on predictions derived from historical values. If markets are efficient long-short type of trading strategy should not be able to provide abnormal returns.

Based on the long-short setting it is possible to further analyze predictable patterns behind ETF premiums. Here methodology of Kreis and Licht (2018) is followed. Based on their study, long-short strategy relies on the mean reversion of ETF premiums and the fact that this reversion occurs mainly in ETF trading price. First, overall mean reversion is examined by taking both ETFs having the highest discounts and ETFs having the highest premiums at time t. For each time t, average premium level for these ETFs is calculated. This is extended to one day before and after time t. It is assumed that premiums and discounts peak at time t and then revert towards value 0. To test whether this reversion occurs mainly in ETF trading price, the average change for ETF trading price and NAV is calculated for the same ETFs for both long and short setting. One-sample t-test is then used to test whether these average changes are significantly different from 0.

4. EMPIRICAL RESULTS

4.1 Mispricing against NAV

Empirical analysis begins with closer examination of ETF premium statistics. Table 4 provides detailed analysis of ETF premiums relative to their NAVs. The results are presented individually for each ETF category. Statistical significance is tested for average premium with one-sample t-test. Hypothesized mean premium is 0. Considering the overall sample, it seems that, on average, ETFs trading in LSE are rather efficiently priced. Average premium relative to NAV is only 0.0464% or 4.64 basis points (bps) which indicates almost non-existent deviation between trading price and NAV. Despite statistical significance the economic magnitude of premium less than 5 bps is not significant. However, a closer inspection reveals differences between different categories of ETFs. Some ETFs tend to trade at significant discount or premium against their NAVs from both statistical and, on some degree, from economic perspective. Examples of such ETF types are high-yield and corporate bonds, Latin American and UK-based equities. Especially high and significant premiums of the UK equity ETFs are rather surprising as domestic equity ETFs have generally been considered most efficient ETFs. It is also notable that some ETFs like US equities or government bonds have basically non-existent average premium at group level.

As stated in previous literature, examination of just average premiums can be misleading. Despite trading at small average premium ETFs can exhibit premium volatility. This may lead to larger deviations from time to time. Typical ETF has premium volatility of 47.59 bps which translates into +93.28 to -93.28 bps range at the 95% confidence level. Considering both positive and negative sides, a confidence band can be calculated and used as a measure for how much premiums can fluctuate around NAV. With this respect our typical ETF would exhibit premium interval of almost 2%. Extreme cases measured with minimum and maximum observations exhibit serious deviation as maximum premium found in 4-year sample is over 17%. Inspection of premium volatility and extreme observations reveals that despite ETFs are trading close to their NAVs on average, significant periodic mispricing can occur.

Table 4: Premiums. Table contains statistics for premiums calculated for each ETF investment style. N of obs. is the number of observations, while N of ETFs is the number of ETFs included within ETF category. Significance of average premium is tested with one sample t-test where hypothesized mean premium is 0. Confidence range and confidence band are calculated based on premium volatility. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Category	N of obs.	N of ETFs	Premium avg. (bps)	Volatility (bps)	95% confidence range	95% confidence band	Min. Premium (bps)	Max. Premium (bps)	Avg. Bid-Ask spread (bps)
All	170923	201	4.64***	47.59	93.28	186.55	-819.99	1753.85	37.37
Equity	121139	143	3.33***	53.51	104.88	209.76	-682.26	1753.84	39.51
EU	23959	27	6.64***	22.96	45.00	90.00	-399.86	881.04	29.04
Japan	6485	7	2.08***	85.80	168.19	336.34	-592.18	639.74	24.17
UK	14850	17	20.34***	26.13	51.21	102.43	-682.26	638.40	45.73
US	20115	28	0.79*	60.01	117.62	235.24	-546.53	1753.85	35.86
Asia	15074	16	-5.97***	72.18	141.47	282.95	-676.47	700.53	42.41
Emerging	9635	11	3.61***	58.18	114.03	228.07	-497.31	722.77	51.43
Latin America	4044	4	-20.11***	75.16	147.31	294.63	-500.95	367.15	39.98
Misc. regions†	6586	7	-0.25	51.07	100.10	200.19	-425.68	1154.57	55.23
World	11440	13	4.17***	34.91	68.42	136.85	-385.45	434.07	42.27
Sector	8951	13	0.32	58.32	114.31	228.61	-614.82	609.44	43.54
Bond	37831	45	9.46***	24.36	47.75	95.49	-819.99	244.00	27.77
Government	17115	19	-1.11***	18.26	35.79	71.58	-217.36	244.00	23.35
Corporate general	13391	17	16.39***	25.06	49.12	98.24	-819.99	208.94	32.45
High-Yield	4292	6	33.28***	27.05	53.02	106.04	-219.32	172.01	42.91
Diversified	3033	3	3.84***	9.96	19.52	39.04	-82.69	108.29	10.50
Miscellaneous	11953	13	2.70***	37.27	73.05	146.10	-322.39	666.32	45.96
Commodity	2854	4	1.28	40.41	79.20	158.41	-226.81	275.31	68.23
Levered	3033	3	0.85	26.74	52.41	104.82	-149.42	666.32	28.79
Property	6066	6	4.29***	40.09	78.58	157.15	-322.39	309.46	44.03

†Misc. region category covers individual country ETFs that do not clearly fall into broader categories. These countries include Australia, South Africa, Russia and Turkey. In contrast, category Emerging equities includes ETFs tracking broader emerging market indices. Notably, categorizing these ETFs differently does not qualitatively alter the results presented in this thesis.

Considering different ETF categories, more exotic equity ETFs such as the Asian equities, have the highest premium volatilities, while the UK and European ETFs have the lowest. This partly highlights the effect of stale NAV issues, as ETFs having underlying assets in time-zones different from the UK are more prone to this issue. It may also explain why the US equity ETFs exhibit high premium volatility from the perspective of European investors. Also, corporate bonds and high-yield bonds exhibit large average premium and quite low premium volatility. This is expected and explained by using bid-quote in NAV calculation as Fulkerson et al. (2014) has described.

In Appendix 3 the sample is divided into two equal-length sub-periods: 2-year periods of 2014-2015 and 2016-2017. The result for these sub-periods are consistent with those for the full-sample period. The same ETF types exhibit larger premium level and volatility in all sample periods being tested. Only the US equity ETFs exhibit large decrease in premium volatility in the period of 2016-2017. This result is partly explained by the difference in minimum and maximum observations between the sample periods. These values during the first period are caused by a newly issued ETF exhibiting higher premium volatility. Despite the overall time period is not long, the results suggest consistency over time.

Table 5 presents the premiums that have been estimated using the model provided by Engle and Sarkar (2006) and Delcoure and Zhong (2007). Controlling staleness of NAVs acts as a robustness check for the results shown in Table 4. If the model performs properly, premium volatilities should decrease especially in cases of international ETFs. This seems to be the effect of the adjustment as premium volatilities decrease to some extent. However, average level of premium is not affected by the adjustment, as the results remains qualitatively similar.

Table 5: The Engle-adjusted premiums. Table contains statistics for premiums calculated with the Engle adjustment for each ETF investment style. N of obs. is the number of observations, while N of ETFs is the number of ETFs included within ETF category. Significance of average premium is tested with one sample t-test where hypothesized mean premium is 0. Confidence range and confidence band are calculated based on premium volatility. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Category	N of obs.	N of ETFs	Premium avg. (bps)	Volatility (bps)	95% confidence range	95% confidence band	Min. Premium (bps)	Max. Premium (bps)	Avg. Bid-Ask spread (bps)
All	170720	201	4.45***	42.41	83.12	166.25	-708.09	1404.59	37.37
Equity	120994	143	3.09***	47.31	92.73	185.46	-708.09	1404.59	39.51
EU	23932	27	6.70***	22.24	43.59	87.18	-393.67	658.56	29.04
Japan	6478	7	0.62	69.88	136.96	273.93	-521.25	638.33	24.17
UK	14833	17	20.11***	25.96	50.88	101.76	-708.09	569.57	45.73
US	20086	28	1.78***	50.93	99.82	199.65	-508.52	1404.59	35.86
Asia	15058	16	-7.36***	62.60	122.74	245.47	-662.83	645.33	42.41
Emerging	9623	11	2.50***	52.45	102.80	205.60	-469.07	691.02	51.43
Latin America	4040	4	-21.04***	71.57	140.28	280.55	-500.77	347.95	39.98
Misc. regions	6579	7	-0.99*	47.51	93.12	186.24	-421.54	944.14	55.23
World	11427	13	4.39***	32.37	63.45	126.89	-377.69	459.99	42.27
Sector	8938	13	0.36	51.12	100.20	200.39	-587.54	483.48	43.54
Bond	37786	45	9.35***	24.01	47.06	94.12	-608.22	233.50	27.77
Government	17096	19	-0.978***	17.85	35.38	70.76	-218.65	233.5	23.35
Corporate general	13374	17	16.24***	24.61	48.24	96.47	-608.22	203.62	32.45
High-Yield	4286	6	32.84***	26.62	52.18	104.35	-209.24	152.91	42.91
Diversified	3030	3	3.86***	9.87	19.35	38.70	-81.89	107.84	10.50
Miscellaneous	11940	13	2.74***	33.47	65.60	131.20	-258.40	646.69	45.96
Commodity	2850	4	1.21*	36.25	71.05	142.10	-166.44	238.09	68.23
Levered	3030	3	0.97**	26.39	51.72	103.45	-154.62	646.69	28.79
Property	6060	6	4.34***	35.17	68.93	137.87	-258.40	293.70	44.03

Adjusting premium for stale NAVs has a different effect on different types of ETFs. Premium volatility of international equity ETFs decreases the most. For bond ETFs the adjustment effect is rather small, which is explained by the choice of equity portfolio used in adjusting the model. Equity-based market portfolio is not the most suitable for bond ETF adjustments. Another effect of the Engle adjustment is decrease in extreme values as it reduces both positive and negative extreme values for most ETFs. According to the results the adjustment model can be considered to correct at least some degree of staleness in NAVs especially with equity ETFs.

Appendix 4 provides additional outlook for premium analysis. In Appendix 4 absolute premiums are used in analysis instead of regular premiums. The results of this type of analysis provide qualitatively similar results regarding mispricing of ETF categories. Also, premiums volatilities do exist in this type of analysis. Similarly, the usage of the Engle-adjusted premiums instead cuts off magnitude in cases of some ETF categories, namely for the Japanese, Asian and US equity ETFs.

4.2 Premium persistence

ETFs can exhibit high premiums due to the premium volatility. However, these premiums may not be long lasting for investors to capture. This leads to the question of premium persistence. The results in Table 6 are derived from AR-models estimated for both unadjusted and the Engle-adjusted premiums. Each ETF is modelled individually. Due to space limitations the results are presented as a summary similarly to the ones obtained by Rompotis (2010). Here number of statistically significant coefficients are presented for each ETF category based on the estimated AR-models. Significant alpha can be inferred as significant average premium while beta coefficients measure persistence. If the first beta is significant, premium persists one trading day. Premium persists two trading days if two betas are significant. The remaining betas follow similar logic.

Table 6: Premium persistence. Table contains summary of regressions of AR-models: $PREM_t = \alpha + \beta_1 PREM_{t-1} + \beta_2 PREM_{t-2} + \beta_3 PREM_{t-3} + \beta_4 PREM_{t-4} + \beta_5 PREM_{t-5} + \varepsilon_t$. Each ETF premium time-series is modelled individually starting by inclusion of just one lag (t-1). Lagged values are then added until at least one beta becomes insignificant. 5% significance level is used here. The results are then summarized at ETF category level. The numbers for significant cases for each regression coefficient are shown in columns. The numbers of significant coefficients calculated with the Engle-adjusted premiums are in parentheses. The results are based on the full sample period of 2014-2017.

Category	N of ETFs	α	β_1	β_2	β_3	β_4	β_5
Equity	143	81 (83)	111 (132)	77 (88)	53 (47)	24 (29)	14 (15)
EU	27	24 (24)	26 (26)	20 (19)	13 (12)	8 (9)	2 (2)
Japan	7	0 (0)	7 (7)	1 (0)	1 (0)	0 (0)	0 (0)
UK	17	14 (14)	17 (16)	14 (15)	13 (13)	12 (12)	10 (10)
US	28	5 (6)	13 (25)	4 (13)	2 (1)	1 (0)	0 (0)
Asia	16	10 (11)	16 (16)	16 (15)	9 (8)	0 (0)	0 (0)
Emerging	11	8 (8)	11 (11)	9 (8)	8 (8)	1 (6)	0 (2)
Latin America	4	4 (4)	3 (4)	2 (2)	2 (2)	2 (2)	2 (1)
Misc. regions	7	4 (4)	7 (7)	6 (6)	3 (2)	0 (0)	0 (0)
World	13	8 (8)	5 (13)	3 (9)	2 (1)	0 (0)	0 (0)
Sector	13	4 (4)	6 (7)	2 (1)	0 (0)	0 (0)	0 (0)
Bonds	45	34 (33)	33 (32)	30 (30)	22 (24)	14 (13)	2 (5)
Government	19	12 (12)	10 (9)	8 (8)	6 (7)	2 (2)	0 (2)
Corporate general	17	14 (14)	15 (15)	15 (15)	13 (13)	10 (9)	2 (3)
High-Yield	6	5 (5)	6 (6)	5 (5)	2 (2)	1 (1)	0 (0)
Diversified	3	3 (2)	2 (2)	2 (2)	1 (2)	1 (1)	0 (0)
Miscellaneous	13	6 (7)	9 (9)	4 (4)	3 (2)	2 (2)	2 (2)
Commodity	4	3 (3)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Levered	3	1 (1)	2 (2)	0 (0)	0 (0)	0 (0)	0 (0)
Property	6	2 (3)	6 (6)	4 (4)	3 (2)	2 (2)	2 (2)
All	201	121 (123)	153 (173)	111 (122)	78 (73)	40 (44)	18 (22)

Based on the overall results we can see that majority of ETFs in the sample has at least one significant beta indicating that premiums persist at least for a trading day. The second lagged premium has explanatory power in cases of 111 ETFs, which counts slightly over half of the sample. Two-day premium persistence is not a major sign of inefficiency but suggests that investors have possibilities to trade against mispricing on a daily level. Some ETFs have premiums with explanatory power even with fifth lag indicating highly persistent premiums.

The most surprising result is high persistence of the UK equity ETF premiums. The result of significant higher average premium and long persistence are against expectations of strong pricing efficiency of domestic ETFs. Also, it should be considered that high persistence does not automatically mean high mispricing. Especially if average level of premium is low. For example, even two-day persistence in premiums of Asian or Latin American equity ETFs can be significant for investors as these ETFs are more prone to higher levels of mispricing based on the results of Table 4. Following the same logic high persistence of the EU equity premiums does not suggest that these ETFs are highly mispriced as the average premium and premium volatility is low. In case of the EU equity ETFs, arbitrage for small premiums may not be interesting to APs. Thus, small premiums can persist. The situation is also seen, for example, with ETFs following government bonds.

The results calculated based on the Engle-adjusted premiums are provided in parentheses. Persistence increases to small extent when the Engle-adjusted premiums are used. The results align with Petäjistö (2017) who found an increase in autocorrelation of ETFs when stale NAVs are controlled for. This is an expected result as these “true premiums” can be considered to last longer when the effect of stale NAVs is controlled for. Change in premium persistence can be seen, for example, in case of the US and World equity ETFs at the first two betas. For the emerging market ETF premiums, the adjustments increase persistence even to fourth and fifth lag in case of some ETFs.

Appendix 5 provides results of persistence analysis done in panel data setting. The usage of panel data and modelling with combined sample seemingly increases premium persistence in case of some ETF categories. However, other important aspect of this type of analysis is a decreasing trend in beta coefficients. Similar trend can be seen in most of the categories. The results indicate that previous day’s premium has the largest effect over today’s premium. This effect decreases dramatically for the second beta suggesting that an investor interested in trading against mispriced ETFs should act rather quickly to capture the identified deviations.

4.3 ETF level determinants behind premium magnitude

Table 7 provides results for ETF level determinants that are assumed to affect mispricing. Here mispricing is measured as absolute premium, which is regressed against variables assumed to have explanatory power. Results are reported for sample containing all ETFs and then dividing ETFs broadly based on underlying assets as presented in Table 4. All models assume fixed effects. Appropriateness for the fixed effect models is tested with the Hausman test and the F-test for no fixed effects. The results of these tests are also reported in Table 7. Also, the models using random effects assumption were estimated but they are omitted on the basis of the results of the Hausman test.

Based on the results, bid-ask spreads of ETFs have statistically significant positive relation to premium magnitude in the all estimated models. As expected, widening bid-ask spread leads to a larger premium magnitude, because coefficients have predicted signs. Since the dependent variable is absolute value of premium, results indicate that overall widening bid-ask spread pushes prices further away from NAV regardless of whether this deviation is negative or positive premium. Other factor having statistically significant effect is volatility of NAV. If ETF is following more volatile assets it can exhibit larger deviation from its NAV. Both size and age were expected to have a negative relationship with absolute premiums. Notably size has significance only with miscellaneous ETFs while age is not significant in any model. Further look towards ETF age reveals that it may have a relation with premium level as suggested by findings of Picotti (2018). ETFs older than 5 years have average premium level of 0.439 bps whereas ETFs younger than 5 years have premium level of 11.13 bps. This is consistent with Picotti (2018) and implies that the age is linked more to premium level rather than premium magnitude.

Table 7: ETF level determinants for premium magnitude. Panel regressions of model: $|PREM|_{it} = \alpha + \beta_{SPDR}SPRD_{it} + \beta_{VOLNAV}\ VOLNAV_{it} + \beta_{SIZE}SIZE_{it} + \beta_{AGE}AGE_{it} + \varepsilon_{it}$. Both time and ETF fixed effects are included in each regression. T-statistics in parentheses are based on heteroscedasticity robust standard errors clustered at ETF level. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Sample		All	Equity	Bond	Misc.
Variable	Predicted sign				
Intercept		0.242***	0.288***	0.116***	0.253***
		(9.61)	(8.97)	(5.40)	(7.56)
Bid – Ask spread	+	0.065***	0.056***	0.074***	0.107***
		(4.32)	(3.62)	(3.86)	(3.57)
NAV volatility	+	0.092***	0.094**	0.035***	0.005
		(2.77)	(2.11)	(3.58)	(0.84)
Size	-	-0.011	-0.015	0.005	-0.010***
		(-1.52)	(-1.51)	(1.10)	(-2.83)
Age	-	-0.007	-0.007	-0.010	-0.008
		(-0.55)	(-0.39)	(-0.60)	(-0.32)
Hausman test		<.0001	<.0001	0.0001	0.0054
F-test for no fixed effects		<.0001	<.0001	<.0001	<.0001
N		169 750	120 269	37 602	11 879
Adj. R2		0.292	0.282	0.523	0.254

Regarding individual significant coefficients the values are at similar levels than in previous studies. Based on rather low values of significant coefficients there is no strong movement in absolute premium when each of the significant variable increases. However, this aligns with results of previous studies, for example, Delcoure and Zhong (2007).

Regarding overall model performance, the inclusion of all ETFs results in the adjusted R-squared of 0.292, which is a bit lower than reported for similar models in previous literature. However, the difference may be explained by the fact that some researchers have also added some form of lagged values of the dependent variable in the model. Regarding already found strong persistence with some ETFs, this may have led to increase in the adjusted R-squareds.

In general, differences between different types of ETFs are rather small. Bid-ask spread and NAV volatility have statistical significance almost regardless of the ETFs used in the model. Notable exception in case of NAV volatility is loss of significance with the sub-sample of miscellaneous ETFs. For bond ETFs, similar variables can explain larger portion of variation. However, in case of the full sample, relatively low adjusted R-squared and significant intercept indicate that much is still left unexplained regarding ETF premium magnitude.

Panel data models estimated in Table 7 should follow similar assumptions than traditional OLS regression models. Initial tests revealed models to exhibit cross-sectional and serial correlation as well as heteroscedasticity. Petersen (2009) and Thompson (2011) have both studied appropriate adjustments in financial studies using panel data and conclude that wrong adjustments may lead to incorrect inferences. Both authors suggest clustering standard errors in case of panel data to avoid incorrect inferences. To overcome these issues, the models were estimated following approaches of Badenhorst (2017), Fulkerson et al. (2014) and Ben-David et al. (2012) as all the authors used rather similar models in their studies. The standard errors for the reported t-statistics are obtained using White's heteroscedasticity robust standard errors which are clustered by cross-sectional variable, in this case by ETFs, following Fulkerson et al. (2014) and Ben-David et al. (2012). Furthermore, Thompson (2011) suggests that the approach of clustering with smaller dimension is more appropriate, especially when N and T in panel have seemingly different sizes. The models were also estimated with double clustered standard errors presented by Petersen (2009) and used by Badenhorst (2017). These unreported models yielded qualitatively consistent results, thereby indicating the robustness of results. Thompson (2011) further discusses that the number of clusters should be sufficient for clustering method to succeed. Miscellaneous ETF category can be argued to have less clusters than optimally wanted. However, if compared to the model reported in Table 7, usage of un-clustered standard errors with miscellaneous ETFs does not qualitatively alter the significance of coefficients when tested. For consistency purposes, the results of the model with clustered standard errors is reported. In addition, residual distribution did not follow normal distribution. However, due to sufficiently large sample sizes in each model, violation of this assumption can be considered as somewhat inconsequential. This conclusion is drawn from central limit theorem (Brooks 2014, 209-211).

Robustness analysis is done by replacing the dependent variables with absolute premiums calculated based on the Engle adjustments. Table 8 presents results for this analysis. Qualitatively results do not change if different dependent variable is used. This is consistent with results of Delcoure and Zhong (2007) who also found similar results by using differently estimated absolute premiums as dependent variables. Also, the results are rather consistent if the sample period is divided into two 2-year sub-samples. The results on this sub-period analysis are not tabulated, but they support the findings presented in Tables 7 and 8 as the same variables hold their significance when the full sample of ETFs is used. Regarding categorical sub-samples some differences can be observed within shorter time periods. NAV volatility loses its significance with equity and bond ETFs during the period of 2014-2015 while bid-ask spread loses its significance during the period of 2016-2017 with equity ETFs. These results could be mainly caused by relatively short sub-periods. Moreover, previous studies covered in reviewed literature report result based on longer sample periods.

Table 8: ETF level determinants for magnitude of the Engle-adjusted premiums. Panel regressions of model: $|PREM|_{it} = \alpha + \beta_{SPDR}SPRD_{it} + \beta_{VOLNAV}\, VOLNAV_{it} + \beta_{SIZE}SIZE_{it} + \beta_{AGE}AGE_{it} + \varepsilon_{it}$. Both time and ETF fixed effects are included in each regression. T-statistics in parentheses are based on heteroscedasticity robust standard errors clustered at ETF level. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Sample		All	Equity	Bond	Misc.
Variable	Predicted sign				
Intercept		0.227***	0.262***	0.103***	0.190***
		(9.86)	(8.72)	(3.81)	(8.86)
Bid – Ask spread	+	0.061***	0.051***	0.078***	0.106***
		(4.06)	(3.26)	(4.10)	(3.61)
NAV volatility	+	0.079**	0.090**	0.034***	0.005
		(2.53)	(2.16)	(3.68)	(0.61)
Size	-	-0.010	-0.013	0.004	-0.004
		(-1.55)	(-1.54)	(0.87)	(-1.38)
Age	-	-0.012	-0.019	0.016	0.012
		(-1.18)	(-1.41)	(1.38)	(0.58)
Hausman test		<.0001	<.0001	0.0002	0.0492
F-test for no fixed effects		<.0001	<.0001	<.0001	<.0001
N		169 748	120 267	37 602	11 879
Adj. R2		0.277	0.258	0.532	0.249

4.4 Market determinants behind premium dispersion

In addition to ETF level determinants, inefficiencies in pricing may occur from market determinants. Table 9 presents results of a set of regressions aiming to explain cross-sectional premium dispersion with market factors. As in previous analysis, the model is first estimated using sample of all ETFs, then with categorical sub-samples and lastly with premium dispersion calculated based on the Engle-adjusted premiums. The results suggest that all explanatory variables can explain premium dispersion. However, significance of the used variables varies between ETF-categories as only FTSE 100 VIX remains significant in each model. Notably, almost all independent variables maintain the predicted signs in each model. Premium dispersion increases when markets exhibit higher volatility and widening TED spread. FTSE All shares index has the opposite effect as dispersion decreases when markets exhibit positive returns.

Table 9: Premium dispersion. Table contains results for model: $PREMSTD_t = \alpha + \beta_{VIX}VIX_t + \beta_{MARKET}MARKET_t + \beta_{TED}TED_t + \varepsilon_t$. T-statistics in parentheses are based on the Newey-West (1987) standard errors with lag length of 6. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Sample		All	Equity	Bond	Misc.
Variable	Predicted sign	F<0.0001	F<0.0001	F<0.0001	F<0.0001
Intercept		0.040	0.043	0.083***	0.057
		(0.75)	(0.67)	(4.14)	(1.44)
FTSE 100 VIX	+	0.023***	0.025***	0.006***	0.017***
		(7.48)	(7.09)	(5.89)	(7.68)
FTSE All Shares returns	-	-0.027**	-0.030**	0.004	-0.013**
		(-2.55)	(-2.42)	(0.59)	(-2.32)
UK TED spread	+	0.219	0.207	0.314***	0.094
		(1.58)	(1.25)	(4.55)	(0.98)
N		1011	1011	1011	1011
Adj. R2		0.269	0.244	0.115	0.217

When comparing results between different sub-samples it seems that FTSE 100 VIX is highly significant in each model indicating strong relationship with premium dispersion regardless of ETF categories used. With FTSE All Shares returns and TED spread there are differences. In case of bond ETFs, performance of the UK stock market does not seem to have effect on premium dispersion. TED spread is not a significant explanatory variable for other than the bond-based ETFs. Regarding adjusted R-squareds, none of the estimated models has high explanatory power. Overall, these differences between the estimated models indicate that there are some degree of differences considering market factors influencing pricing inefficiency between different ETF categories.

The results shown in Table 9 are subject to assumptions behind the OLS regression. Diagnostic tests revealed both autocorrelation and heteroscedasticity in the models. The approach of Petäjistö (2017) is followed and the Newey-West (1987) standard errors are used to correct effects of the previously-mentioned issues. In lag selection Brooks (2014, 201) was followed resulting in the usage of 6 lags. All the models reject residual normality. However, sample size remains still sufficiently large to counter these issues. As mentioned in Data and Methodology section also outliers in data were treated in similar manner than in previous studies.

Robustness analysis with the Engle-adjusted premiums confirms some of the effects presented previously. The results of similar analysis based on the Engle-adjusted premiums are presented in Table 10. Based on these results, it seems that ETFs trading in LSE, at least to some extent, follow similar fundamentals than ETFs trading in the US where these factors were originally tested. Especially the effect of FTSE 100 VIX index is supported as variable retains significance in the same ETF categories as in Table 9. Notably, UK TED spread becomes significant for sample using all ETFs while the FTSE All Shares index loses its significance in all categories except with the bond ETFs. Regarding the bond ETF category, significance in FTSE All shares index is seen only after the Engle adjustment. This seems to be a consequence of using FTSE All Shares index as an explanatory variable in the stale NAV adjustment model. Regarding both equity and miscellaneous ETF categories, the stale NAV adjustment method seems to already control the effects of FTSE All Shares index resulting non-significant coefficients. For the bond ETF category, the stale NAV adjustment method seems to create relation to FTSE All Shares index. Moreover, between the models using unadjusted and the Engle-adjusted premiums, only the dependent variable is altered. This further highlights the fact that the stale NAV adjustment with equity-based market index may not be the most proper approach for bond ETFs. The results further imply, that the possible staleness in NAVs of bond ETFs could be more likely caused by illiquidity of underlying assets rather than time-zone mismatches. This type of reasoning regarding differences between international equity and bond ETFs is presented, for example, by Petäjistö (2017).

The explanatory power of the models decreases in every case when premiums are adjusted, except with the bond ETFs. This could be, at least on some extent, caused by the previously-described effects of FTSE All Shares index. Notably, with the models having individual ETF characteristics as independent variables, the difference between the premiums used as the dependent variable did not drastically affect the explanatory power of the models.

Table 10: Premium dispersion based on the Engle-adjusted premiums. Table contains results for model: $PREMSTD_t = \alpha + \beta_{VIX}VIX_t + \beta_{MARKET}MARKET_t + \beta_{TED}TED_t + \varepsilon_t$. T-statistics in parentheses are based on the Newey-West (1987) standard errors with lag length of 6. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Sample		All	Equity	Bond	Misc.
Variable	Predicted sign	F<0.0001	F<0.0001	F<0.0001	F<0.0001
Intercept		0.096**	0.105**	0.087***	0.084***
		(2.30)	(2.11)	(4.64)	(2.86)
FTSE 100 VIX	+	0.016***	0.018***	0.005***	0.012***
		(7.37)	(6.78)	(6.05)	(7.97)
FTSE All Shares returns	-	-0.013	-0.013	-0.012**	-0.008
		(-1.46)	(-1.31)	(-2.39)	(-1.45)
UK TED spread	+	0.247**	0.238	0.306***	0.114
		(1.97)	(1.58)	(4.66)	(1.42)
N		1010	1010	1010	1010
Adj. R2		0.195	0.170	0.147	0.157

In addition, the models are estimated by dividing sample into two 2-year sub-periods. These results are unreported but qualitatively only FTSE VIX 100 holds its significance constantly. Figure 3 illustrates the relationship of the index and premium dispersion of all ETFs in the sample. Otherwise, the results are more mixed regarding TED spread and the FTSE All shares index. Especially level of TED spread changes between 2014-2015 and 2016-2017, which indicates the reason behind the differences between the sub-periods. Moreover, TED spread is significant during the period of 2014-2015 with all sub-samples except bond ETFs (also for bonds if 10% significance level is used). Increase in TED is seen after Brexit-vote which leads to loss of significance for sub-samples using all and equity ETFs in the latter sub-period, while with bond ETFs TED spread becomes highly significant. In previous peer-group studies the models were estimated on the basis of significantly longer periods. For example, Petäjistö (2017) used the period of 2007-2014 with inclusion of the financial crisis. Since the used sub-periods in this thesis are rather short, they can influence the obtained results to some extent. However, the results of the sub-period analysis suggest that the FTSE 100 VIX index is the best benchmark regarding ETF pricing efficiency in LSE. The relationships with other variables are not constant.

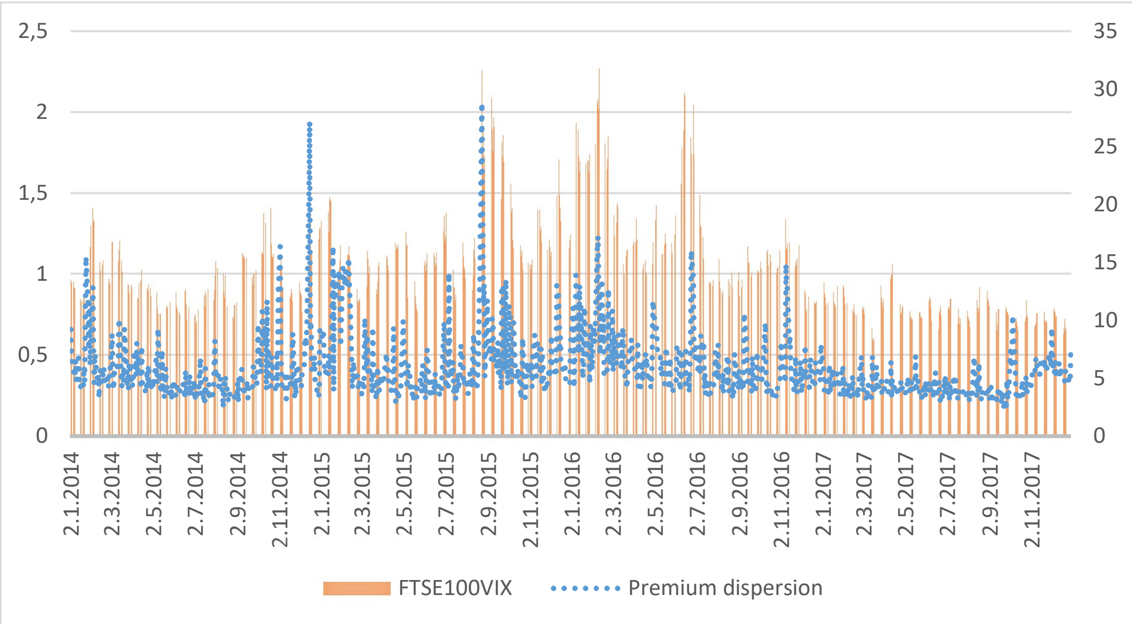


Figure 3: Premium dispersion and FTSE 100 VIX

4.5 Long-short trading against mispricing

Lastly ETF mispricing is examined in terms of active trading strategy. The selected Filter-strategy is implemented with long-short positions based on estimated premiums. The Filter-strategy takes positions on ETF if premium is above or the below selected filter. Table 11 provides summary of results regarding strategies and their performance for different samples of ETFs. The usage of active trading without transaction costs yield qualitatively similar results than reported in previous literature. For example, Charteris (2013), Jares and Lavin (2004), Petäjistö (2017) and Kreis and Licht (2018) all found high cumulative returns before transaction costs.

As mentioned in the description of the trading strategy, the same closing prices are used for signaling and trading. Also, premiums based on official NAVs are available only after the close of LSE which creates potential look-ahead bias for ETFs having their underlying assets traded after the close of LSE. In order to increase robustness of the results the sample is divided into sub-samples A, B and C. Samples A and B can be considered to be free of look-ahead bias as they contain only ETFs that have their underlying assets closing earlier or around the same time as LSE. As for additional unreported robustness check the Sample A was further narrowed to include only UK equity ETFs. The results remained qualitatively consistent with those reported in Table 11 despite decrease in return potential. However, this seems to be mainly caused by lower number of potential ETFs as active trading days are already few in number for Sample A.

Following Simon and Sternberg (2005), an upper limit of 200 bps is introduced for filters regarding Sample C using all ETFs. These large premiums in case of ETFs having assets trading in US time-zones have likely been influenced by the bias. As shown by Table 11, the results do vary between different samples. Sample C has very high cumulative return compared to the two other samples A and B. Further inspection of Sample C reveals that ETFs having their underlying assets traded after the close of LSE have seemingly better return predictability based on the trading signals. Also, high returns are mainly driven by a couple of ETFs having volatile underlying assets and thus, larger return potential (for example Latin American ETFs and ETFs linked to gold producing). If the five most traded ETFs, whose NAVs can be considered biased, are removed from the Sample C, results change dramatically. These results are reported in Appendix 6.

Table 11: Comparison of trading portfolios. Filter-strategy pursues excess return by taking positions against mispriced ETFs. Filter-strategies are tested with different filters to provide estimates of their profit potential, and in addition, to reveal their sensitivity to transaction costs. Lagged strategies follow one-day waiting rule after signal. Buy-Hold consists of equal investment to each ETF in the given sample. FTSE All shares measures return of investment to local market index. The results are based on the full sample period of 2014-2017.

Sample A: UK and EU equities (44 ETFs)	Filter 60	Filter 60 (t-costs)	Filter 80 (t-costs)	Filter 80 lagged	Filter 80 lagged (t-costs)	Buy-Hold	FTSE All shares
Total cumulative return	190.15%	43.70%	35.99%	40.98%	17.03%	26.70%	30.74%
Average daily return	0.108%	0.038%	0.031%	0.035%	0.017%	0.027%	0.030%
Daily standard deviation	0.662%	0.646%	0.436%	0.499%	0.489%	0.833%	0.835%
Max	8.305%	8.089%	7.926%	8.889%	8.671%	3.179%	3.413%
Min	-2.178%	-2.374%	-2.782%	-3.326%	-3.520%	-3.978%	-4.634%
Sharpe ratio	2.853	0.887	1.101	1.085	0.472	0.433	0.495
Filter (+/- bps)	60	60	80	80	80	-	-
Active trading days	351	351	87	93	93	-	-
Transaction cost per roundtrip	-	0.2%	0.35%	-	0.2%	-	-
Sample B: UK, EU, Asia, Japan and Misc. regions (74 ETFs)	**Filter 50**	**Filter 50 (t-costs)**	**Filter 150 (t-costs)**	**Filter 65 lagged**	**Filter 150 lagged (t-costs)**	**Buy-Hold**	
Total cumulative return	55.93%	-40.74%	15.78%	15.88%	-10.33%	33.90%	
Average daily return	0.048%	-0.048%	0.018%	0.020%	-0.007%	0.032%	
Daily standard deviation	0.852%	0.852%	0.864%	1.023%	0.859%	0.844%	
Max	4.513%	4.408%	6.688%	9.013%	8.904%	4.337%	
Min	-4.068%	-4.164%	-7.335%	-6.473%	-6.718%	-5.024%	
Sharpe ratio	0.838	-0.017	0.247	0.210	-0.004	0.537	
Filter (+/-bps)	50	50	150	65	150	-	
Active trading days	967	967	280	837	276	-	
Transaction cost per roundtrip	-	0.1%	0.1%	-	0.1%	-	
Sample C: All (201 ETFs)	**Filter 125-200**	**Filter 125-200 (t-costs)**	**Filter 125-200 (t-costs)**	**Filter 150 lagged**	**Filter 150 lagged (t-costs)**	**Buy-Hold**	
Total cumulative return	1751.35%	227.04%	37.18%	44.83%	-47.93%	33.75%	
Average daily return	0.297%	0.124%	0.038%	0.045%	-0.056%	0.031%	
Daily standard deviation	1.218%	1.192%	1.186%	1.280%	1.277%	0.649%	
Max	10.653%	10.321%	10.155%	8.459%	8.242%	2.724%	
Min	-5.811%	-6.094%	-6.235%	-14.831%	-15.001%	-3.612%	
Sharpe ratio	5.498	1.793	0.417	0.459	-0.031	0.696	
Filter (+/-)	125-200	125-200	125-200	150	150	-	
Active trading days	577	577	577	511	511	-	
Transaction cost per roundtrip	-	0.3%	0.45%	-	0.2%	-	

Active trading generates transaction costs which cannot be ignored as trading occurs on daily frequency. Since the strategy trades daily, the results before any costs are to be considered hypothetical. According to Gastineau (2011), the costs in cases of ETFs are mainly associated with bid-ask spread, as especially large institutional investors can push fixed costs almost to zero. As shown by the previous results, higher bid-ask spreads are associated with higher premiums, from which the strategy aims to profit. However, even smaller premiums may seem attractive before any costs. Furthermore, the inclusion of transaction costs provides more realistic estimate of the performance as returns are heavily affected by their inclusion.

Notably, the strategies can be optimized after inclusion of costs. Figure 4 illustrates cumulative returns of Sample A. It is evident how selection of the filter and potential costs affect the results. Focusing on higher premiums enhances the ability to cope with trading costs for Samples A and B, but significantly decreases the number of trading days. This is in line with Jares and Lavin (2004) who report the number of active days to be only around 10% when strategy is optimized after transactions costs. Furthermore, only Samples A and C can tolerate some degree of transactions costs and outperform their buy-and-hold alternatives. However, their after-transaction-cost performance is not high enough to justify implementation of this type of strategy. Moreover, the average bid-ask spread of the whole sample is 0.37% while some equity ETFs have even higher average spreads. None of the reported strategies cannot absorb such level of costs. This holds even for Sample C with possible bias. Despite potential predictability based on the premiums, daily returns are too low to overcome effects of bid-ask spreads. For Sample B, profit potential is very limited, and inclusion of costs quickly erases returns. Further inspection reveals this to be partly caused by the Asian and Japanese ETFs as they take numerous positions which are unprofitable even for over half of the trades for some ETFs. This is shown by also the results without costs if compared to the other samples. The effect is also evident in Appendix 6 as the Asian-based ETFs accumulate increasing amount of positions when the previously-mentioned biased ETFs are controlled for.

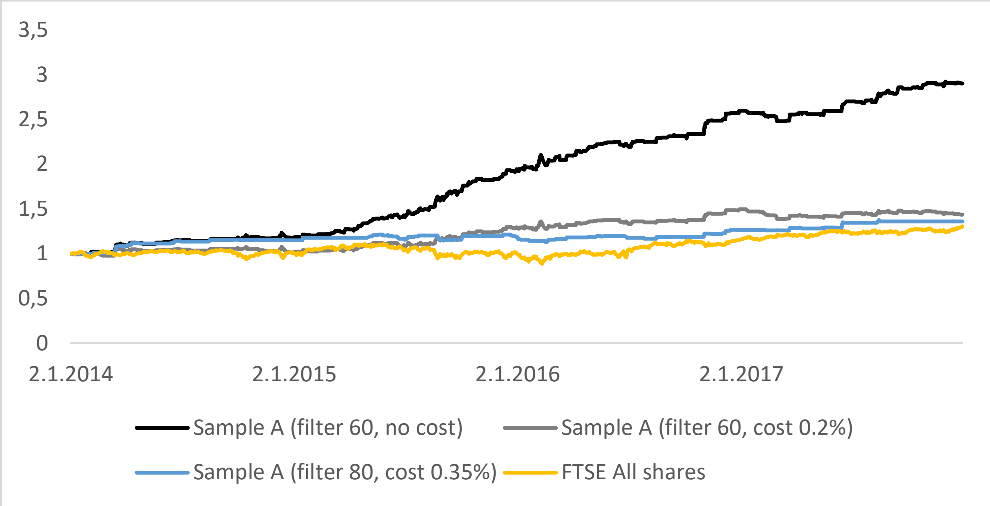


Figure 4: Cumulative returns of FTSE All shares index and Filter-strategies of Sample A

Table 12 presents the results when performance of active trading is benchmarked against the CAPM. Following Petäjistö (2017) and Kreis and Licht (2018), it is tested whether the trading strategies could produce a significantly positive alpha. The results indicate a statistically significant alpha mostly for the strategies without transaction costs. The levels of the alphas are lower than reported by Kreis and Licht (2018) as they report 0.96% before transaction costs and 0.27% after transaction costs when trading with European-based ETFs listed in XETRA. Based on the results of this thesis, the inclusion of transaction costs leads to insignificant alphas. Similar results are documented by Kreis and Licht (2018) if the years of financial crisis are removed from their sample. As is typical for the long-short strategy, their market betas and adjusted R-squareds are rather low. Moreover, the results from active trading are independent from overall performance of the employed benchmarks. This is also partly explained by the days without any trading signal. Every Filter-strategy is employed for higher levels of premiums and thus, most of the estimated strategies have several days when no trading occurred regardless of the samples employed. Qualitatively selection of different benchmark does not alter the results despite exact alphas and betas vary.

Table 12: Performance from the CAPM. Performance of active strategies is examined within framework of the CAPM: $Rp_t - Rf_t = \alpha + \beta(Rm_t - Rf_t) + \varepsilon_t$. Buy-and-hold strategies for each sample reported in Table 11 and market portfolios obtained from data library of Kenneth French are employed as Rm_t. Market portfolio returns of French are provided originally based on US dollars. They are adjusted for currency effects by using daily USD/GBP spot rate obtained from Bank of England database. UK 3-month T-bill is used as Rf_t. T-statistics in parentheses are calculated from heteroscedasticity robust standard errors. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Strategy			B-H benchmarks				Market portfolio benchmarks			
Sample A: UK and EU equities (44 ETFs)	N of trading days	N of active trading days	Alpha (%)	Beta	Adj. R2	Benchmark	Alpha (%)	Beta	Adj. R2	Benchmark
Sample A (Filter 60, no costs)	1010	351	0.113***	-0.259***	0.105	B-H (Sample A)	0.116***	-0.240***	0.097	European
			(5.77)	(-5.74)			(5.85)	(-6.14)		
Sample A (Filter 60, t-cost 0.2%)	1010	351	0.043**	-0.255***	0.107	B-H (Sample A)	0.046**	-0.240***	0.102	European
			(2.27)	(-5.79)			(2.40)	(-6.28)		
Sample A (Filter 80, t-cost 0.35%)	1010	87	0.030**	-0.014	-0.000	B-H (Sample A)	0.031**	-0.021	0.001	European
			(2.27)	(-0.46)			(2.32)	(-0.81)		
Sample B: UK, EU, Asia, Japan and Misc. regions (74 ETFs)										
Sample B (Filter 50, no costs)	1010	967	0.045*	0.039	0.001	B-H (Sample B)	0.047*	-0.008	-0.001	Global ex US
			(1.68)	(0.60)			(1.73)	(-0.13)		
Sample B (Filter 150, t-cost 0.1%)	1010	280	0.010	0.220***	0.045	B-H (Sample B)	0.010	0.165**	0.019	Global ex US
			(0.38)	(2.90)			(0.38)	(2.20)		
Sample C: All (201 ETFs)										
Sample C (Filter 125-200, no costs)	1010	577	0.288***	0.266***	0.019	B-H (Sample C)	0.287***	0.173**	0.011	Global
			(7.61)	(3.00)			(7.45)	(2.23)		
Sample C (Filter 125-200, t-cost 0.45%)	1010	577	0.029	0.265***	0.020	B-H (Sample C)	0.028	0.173**	0.011	Global
			(0.79)	(3.06)			(0.76)	(2.27)		
Sample C (Filter 150, lagged, no costs)	1009	511	0.031	0.431***	0.043	B-H (Sample C)	0.030	0.267***	0.024	Global
			(0.78)	(3.90)			(0.74)	(3.08)		

4.5.1 Predictable patterns and premiums

To examine the behavior of ETF trading prices and NAVs around possible signals, the methodology employed by Kreis and Licht (2018) is followed. Active trading on ETF mispricing has two key requirements: mean reversion of ETF premiums and the fact that this reversion occurs on trading prices of the ETFs. Based on the results of active trading, these requirements are tested following the approach of the authors to further improve understanding of predictable patterns of ETFs used in this thesis.

Figure 5 presents average mean reversion patterns of premiums. For ETFs having the long signal at time t and ETFs having the short signal at time t, average premium level is calculated. This is done for time t, as well as for a day before and after the time t. These are then averaged over time to provide estimates for the full sample period. The estimation is conducted based on the Filter-strategies without transaction costs. Based on the graphs, it is evident that mean reversion occurs on both long and short settings. For example, in the case of Sample C including all ETFs, the average premium on long positions on day t-1 is around -50 bps. Between t-1 and t, the average premium decreases to -150 bps until it reverts to the original level. This explains why the estimated lagged strategies could not outperform against the strategies using more recent information. It also evident, that the strategies applied for Samples A and B are less mean-reverting regardless of the fact that mean reversion occurs in all samples.

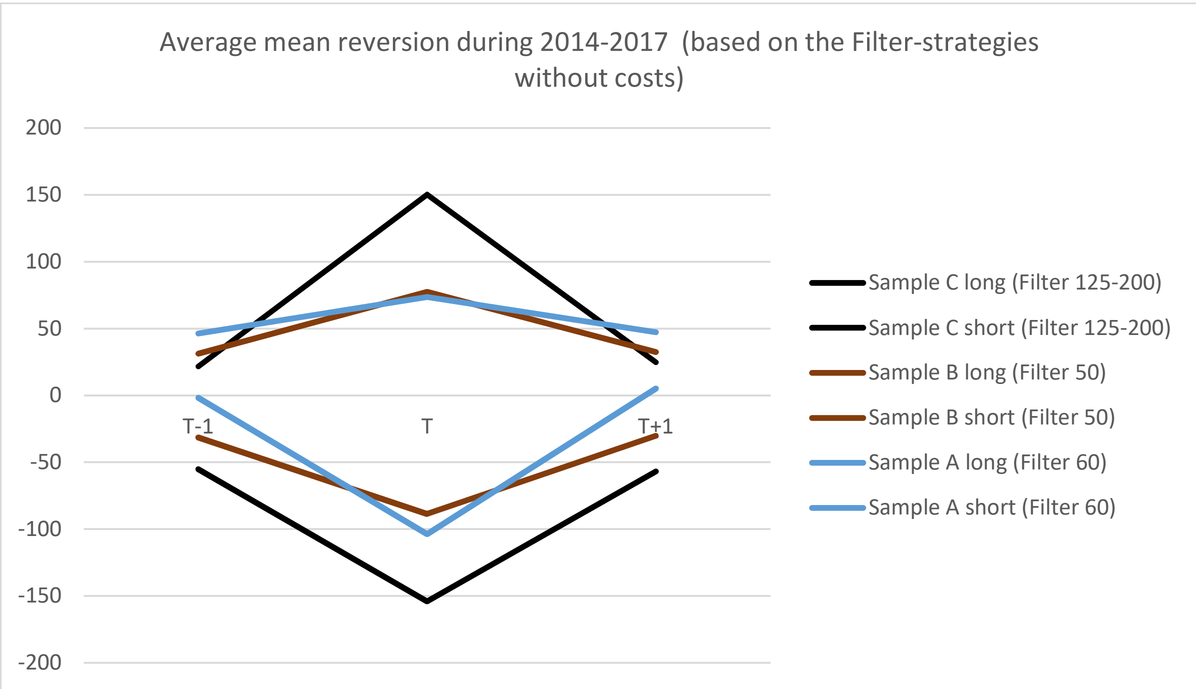


Figure 5: Average mean reversion process during the sample period of 2014-2017

As the mean reversion occurs it can be tested whether the adjustment between days t and t+1 occurs in ETF trading prices. Table 13 provides the results of such analysis. Average change in ETF trading price and NAV is calculated for the same ETFs and positions used in Filter-strategies without transaction costs. Statistical significance is assessed on the basis of one-sample t-test to test whether average change is different from 0. The results indicate that when mispricing occurs both trading prices and NAVs adjust. In the case of long position, trading price increases, and simultaneously, NAV decreases. However, the significance of these changes varies across the samples, while price movements in Sample C are the most significant. These results indicate that there are predictable patterns regarding both ETF trading price and NAV around large premiums. However, as mentioned in the description of trading strategy, investors may be able to capture this movement only partially due to the fact that information and trading possibilities might become available during different times of a day. An additional robustness analysis for Sample C is provided in Appendix 6 with reduced sample. Notably average ETF price movements decrease when the most biased ETFs are left out from the sample. This partly explains a dramatic decrease in returns of trading strategy if these ETFs are removed from the sample.

The trading strategies used in this thesis utilize only ETFs, not underlying assets. Furthermore, the adjustment of ETF prices is needed in this setting. Compared to Kreis and Licht (2018), they reported the adjustment occurring mainly in ETF prices indicating sufficient setting for the trading strategy. Considering Samples A and B, the results in Table 13 offer some additional information regarding estimated performance of the trading strategies. For Sample A, the changes mainly occur in ETF trading price, whereas for Sample B, the changes occur mainly in NAV. This partly explains the poor performance of the Filter-strategy applied on Sample B. Notably, the results in Table 13 suggest that hypothetically investors might be able to enhance returns by taking opposite positions in constituents of ETF and NAV.

The results of both Figure 5 and Table 13 are sensitive to the ETFs included in the analysis. In an unreported analysis, ETFs are assigned similarly than in Kreis and Licht (2018). For each day, ETFs with the highest and the lowest premium are assigned into appropriate positions. If the number of ETFs increases (ETFs with lesser premiums are included), the mean reversion is weaker. Also, both ETF and NAV returns fade towards 0. This suggests that the mean reversion and larger changes in ETF prices and NAVs are associated with higher premiums.

Table 13: Daily changes in ETF price and NAV. ETFs are assigned based on trading signals on day t. Filter-strategies without transaction costs reported in Table 11 are followed. Long position calculates average changes in trading price and NAV between t and t+1 for the ETFs having signal for long positions. Short position calculates the same for the ETFs having signal for short positions. Reported long and short returns are unweighted in contrast to their combined total weighted contributions regarding Filter-strategies. This is seen also in different number of days for each position. One sample t-test, where hypothesized mean is 0, is employed to test the significance of average changes. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Sample A: UK and EU equities (44 ETFS)	Long position	Short position
Return ETF	0.0075***	-0.0027***
Return NAV	-0.0036*	0.0001
N of days	47	316
Sample B: UK, EU, Asia, Japan and Misc.region equities (74 ETFs)		
Return ETF	0.0013**	-0.0006*
Return NAV	-0.0046***	0.0040***
N of days	616	850
Sample C: All (201 ETFs)		
Return ETF	0.0051***	-0.0056***
Return NAV	-0.0045***	0.0072***
N of days	373	284

4.6 Limitations of the results

There are some notable limitations regarding the presented results. First, the robustness analysis is mainly done by adjusting premiums with the model proposed by Engle and Sarkar (2006). The adjustment model relies on several assumptions on the pricing process, which highlights the importance of the selection of independent variables being included in the model. In the US there has been empirical evidence of ETFs following S&P 500 during non-synchronous trading hours. In Europe, similar evidence is not documented. In this thesis, FTSE All shares index is used as an independent variable for premium adjustment as it can be viewed as the market index mostly synchronized with LSE. Following approach of Picotti (2018) who used the US market portfolio in similar manner for sample of multiple different ETF types, the corresponding adjustment was made for all ETF classes. Considering equity ETFs, the effect is similar as in Delcoure and Zhong (2007). For bond ETFs, the approach did not have notable effects. However, this may be the case, because of premiums associated with bond ETFs may be more related to illiquidity of underlying assets. Also, for bond ETFs the adjustment could be more appropriate to implement with synchronous bond index. However, confirming correctness of the approach would require additional research and could be one of the topics to study in the European ETF markets. For this reason, also sub-sampling based on time period was used as an additional robustness check.

The sample of ETFs may also possess some limiting capabilities. Despite the sample being sufficiently large compared to previous studies, it still includes only a fraction of ETFs traded in LSE. This, to some extent, deteriorates the generalizability of the obtained results. However, the degree of diversification within different ETFs included in the sample is sufficient, even though, equity ETFs dominate over other ETF types. On the other hand, the majority of ETFs traded in different exchanges are still equity-based ETFs despite the emergence of more exotic categories.

Regarding the tested trading strategies, the availability of intraday data would be fundamental to conduct robustness analysis, especially for ETFs having their underlying assets traded in different time-zones. As mentioned, the assumption to use same day's closing price and NAV for signals affects the results. This also introduces a look-ahead bias for ETFs having their underlying assets trading after the close of LSE. Notably the Engle adjustment cannot be confidently used as a robustness analysis in this setting as the model also includes information beyond the close of LSE. Also, the number of ETFs in the sample may not have enough trading volume for an investor to take all desired positions, or at least, to enter these positions without affecting the trading prices. Notable, APs should act as liquidity providers for less traded ETFs. Due to these facts and in spite of the previously-described actions to address these issues, the results presented regarding the trading should be viewed as an upper boundary of return potential of active trading strategies aiming to benefit from ETF mispricing. Also, the examined time period does not include any sub-period that could clearly be deemed bearish. Based on the theoretical framework, the inclusion of such period would have been desirable as ETF mispricing is documented to be magnified in bearish conditions. From the viewpoint of trading strategy, this would have been particularly interesting as this would likely enhance profit potential to some extent when trading costs are included.

5. CONCLUSIONS

The objective of this thesis is to discover current state of ETF pricing efficiency in LSE with a variety of different methods proposed by current ETF literature. The sample data consists of 201 ETFs with daily data for the time period from 2.1.2014 to 29.12.2017. Overall, the results are consistent with findings of previous researchers.

The results regarding detected premiums suggest that ETFs trading in LSE are, on average, priced close to their NAVs. However, ETFs traded in LSE are no different from ETFs literature covered in theoretical framework. Evidence of ETF trading prices deviating from NAVs from time to time is found. Thus, ETF pricing inefficiencies exist also in LSE. Most of the average premiums were not economically significant despite their statistical significance. Despite average premium level not indicating mispricing in case of most funds, premium volatility and outlier observations indicate that large premiums do exist across different ETF categories. However, premium volatilities and differences between ETFs are smaller than in Petäjistö's (2017) research. This may be explained by the inclusion of the financial crisis in Petäjistö's (2017) sample period, and by the fact that ETF markets may have generally become more efficient due to the exponential growth in investor's interest, as well as in capital invested. Data used in this thesis is rather recent in this context.

Furthermore, premiums of the UK equity ETFs deviate from the results of previous studies. Most researchers argue that domestic equity ETFs should be the most efficiently-priced ETFs due to their higher liquidity and exemption of stale NAV issues. Especially in the US market this has been evident, as shown, for example, by Hilliard (2014). This is in contradiction with the premium of 20 bps documented for the UK equity ETFs which are domestic ETFs from perspective of this research. 20 bps premium is also notable when comparing it to the premium level of the EU equity ETFs which should represent quite similar ETF category.

Finding evidence of mispricing included examinations on the persistence of the detected premiums. This was tested with AR-models. These results indicate that around half of the

ETFs in the sample had a premium persistence for at least two days. According to Charteris (2013), two-day persistence does not indicate severe pricing inefficiency. However, it is persistent enough for investors trying to exploit mispricing with active trading strategies. The premiums of day t-1 have the strongest effect on today's premiums, indicating that most of the mispricing occurs between first and second day. The findings regarding premium persistence lead to conclusion that premiums persist to some degree regardless of ETF category.

Regarding determinants behind mispricing, the results of different regression analyses suggest that overall market conditions and some ETF characteristics have influence on premium magnitude. Such influence is documented in the case of bid-ask spread and volatility of the underlying assets. Bid-ask spread is argued by many authors to be significant ETF level barrier for APs to do ETF arbitrage as wider spread increases costs of arbitrage for APs. These results strongly align with the ones presented by Delcoure and Zhong (2007) and Ben-David et al. (2012), who also found significance for bid-ask spreads. In this thesis, the significance was also found for volatility of underlying assets. ETFs having more volatile assets as their underlying are more prone to mispricing. This is in line with the findings of international ETFs, especially Asian or Indian based ETFs having higher premiums. The results regarding age and size of ETFs did not provide additional evidence on their relation to the premium magnitude. Differences of results regarding previous studies may be due to the differences in methods and data: Picotti (2018) analyzed ETF age with different methodology and objective. However, the results of this thesis support his view that ETF's age is linked more to premium level than the magnitude.

Overall market conditions do matter for ETF pricing efficiency. The results indicate that similar measures of overall market conditions can be used to explain ETF pricing efficiency also outside the US markets. With this respect, the results align with those of Petäjistö (2017) and Ben-David et al. (2012) in the US markets. Market conditions can be argued to have an effect on AP's ability to do ETF arbitrage with LSE traded ETFs. When arbitrage capital is scarcely available, cross-sectional dispersion of ETF premiums increases. Especially there seems to be clear evidence of the relation to the FTSE 100 VIX index. Also, UK TED spread and FTSE All shares index have some effect on mispricing, but this effect is partly dependent on the ETF-categories under analysis and it varies between shorter sub-periods. In the case of UK TED spread, for example, the results differ before and after Brexit-vote. As most of the APs are global players they may seek arbitrage capital elsewhere after economic situation in the UK altered compared to the rest of the EU area. From economic perspective, ETF investors should consider potential mispricing especially during market distress also in LSE.

Active trading against ETF mispricing serves as convenient summary of the existence of phenomenon. Active trading yields qualitatively similar results than documented in previous studies. Traders reacting quickly on pricing inefficiencies may be able to profit on the matter with both economical and statistical respects if trading costs are not considered. Despite substantial returns, the level of trading costs highly affects the profitability of trading strategies. Thus, potential profits are likely unachievable for investors trading with actual bid-ask spreads. This is in line with findings of Kreis and Licht (2018). However, further analysis on mean reversion process suggests that predictable patterns do exist. Consistent with Ben-David et al. (2012), mean reversion occurs in both ETF trading prices and NAVs, the latter of which was not documented by Kreis and Licht (2018). This type of adjustment toward a new equilibrium decreases profit potential of the strategies that aim to exploit pricing inefficiencies by trading purely ETFs. However, even retail investors may use this information when trading ETFs to enhance their returns. For example, investors could enter long positions when ETF is trading at substantial discount while avoiding to enter into these positions while ETF is trading at substantial premium. Thus, from economical perspective, the results are relevant for investors despite systematical exploitation of mispricing may not be possible.

The last objective of this thesis was to examine whether there are differences regarding mispricing between different types of ETFs. Also with this respect, the overall results are qualitatively similar to previous studies. Generally, mispricing is especially associated with international ETFs. The results are consistent with the earlier literature, as the largest premium volatilities were found in cases of international equity ETFs. This is partly explained by the time-zone mismatches of LSE and trading hours of assets underlying the international ETFs. Regarding premium persistence, clear differences between ETF categories exist. Some ETFs, (for example, commodity ETFs) exhibit almost non-persistent premiums. In contradiction, the Asian equity ETFs exhibit both high premium volatility and daily persistence. The results on the determinants behind mispricing identify small differences between broad ETF categories. Overall, similar characteristics and market conditions can be considered to affect ETF mispricing across different ETF categories despite that the magnitude of the effects may be different depending on the type of ETF under interest. Lastly, there were notable differences in the results between different sub-samples of ETFs regarding the returns of trading strategy and examined mean reversion patterns. Furthermore, the results suggest that the price adjustments in ETF trading price and NAV after large premiums differ between different ETF categories.

As for additional contribution, this thesis also tests the impact of using stale NAV adjustment for ETFs trading in European time-zones. The adjustment method proposed by Engle and Sarkar (2006) for the US domiciled ETFs has expected effects on international equity ETFs examined in this thesis. The reduction of premium volatility due to stale NAV adjustments highlights the need for using adjustment techniques when dealing with international ETFs. As price discovery of European domiciled ETFs is less studied and time-zone mismatches are different from the US markets, the findings with the Engle-adjusted premiums should be treated preliminary only.

Regarding the future, expansion of new ETFs offers a wide range of topics for researchers to address. Especially, research should be done to refine the methodology to adjust for stale NAVs in other countries than US. Many studies highlight importance of adjusting premiums when studying ETF mispricing. One answer might be the usage of iNAVs, especially with US-based ETFs, but historical data on the matter could be unavailable. Also, iNAVs are partly subject to similar staleness issues in cases of the Asian ETFs. Notably there is no research done on stale NAV adjustment using the European ETFs. Also, future studies might include high-frequency trading with intraday data. With this type of data it could be possible to develop and truly simulate a trading strategy to exploit ETF mispricing. Some studies focusing on intraday trading are already available indicating interest towards this area.

REFERENCES

Ackert, L. F. and Tian, Y. (2008). Arbitrage, Liquidity, and the Valuation of Exchange Traded Funds. Financial Markets, Institutions & Instruments, 17(5), pp. 331-362.

Badenhorst, W. M. (2017). Premiums and discounts of exchanged-traded funds. South African Journal of Accounting Research, 31(3), pp. 212-222.

Baltagi, B. H. (2005) Econometric analysis of panel data. 3rd edition. John Wiley & Sons.

Ben-David, I., Franzoni, F. and Moussawi, R. (2012). ETFs, Arbitrage, and Contagion. National Centre of Competence in Research. Working paper No. 793.

Blackrock. (2017). A primer on ETF primary trading and the role of authorized participants. [online]. Available at: https://www.blackrock.com/corporate/en-at/literature/whitepaper/viewpoint-etf-primary-trading-role-of-authorized-participants-march-2017.pdf. [Accessed 10.2.2018]

Bodie, Z., Kane, A. and Marcus, A. (2005). Investments. 6th ed. McGraw-Hill.

Brooks, C. (2014). Introductory Econometrics for Finance. 3rd ed. Cambridge University Press.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. Journal of Finance, *52*(1), pp. 57-82.

Charteris, A. (2013). The price efficiency of South African exchange traded funds. Investment Analysts Journal, 42(78), pp. 1-11.

Charupat, N. and Miu, P. (2013). Recent developments in exchange-traded fund literature. Managerial Finance, 39(5), pp. 427-443.

Chelley-Steeley, P. and Park, K. (2011). Intraday patterns in London listed Exchange Traded Funds. International Review of Financial Analysis, 20(5), pp. 244-251.

Davidson, L. (2013). Examining the exchange-traded nature of exchange-traded funds. Morningstar ETF Research. [White paper]. Available at: http://corporate.morningstar.com/us/documents/MethodologyDocuments/ResearchPapers/exchangetradednature.pdf. [Accessed 8.3.2018]

Delcoure, N. and Zhong, M. (2007). On the premiums of iShares. Journal of Empirical Finance, 14(2), pp. 168-195.

Deville, L. (2008) Exchange Traded Funds: History, Trading and Research. In: C. Zopounidis, M. Doumpos, P. Pardalos. Handbook of Financial Engineering, Springer, pp.1-37.

Elton, E., Gruber, M., Comer, G. and Li, K. (2002). Spiders: Where Are the Bugs? The Journal of Business, 75(3), pp. 453-472

Elton, E. (2003). Modern portfolio theory and investment analysis. 6th ed. Wiley.

Engle, R. and Sarkar, D. (2006). Premiums-Discount and Exchange Traded Funds. Journal of Derivatives, 13 (4), pp. 27-45,6.

Ernst & Young. (2017). Global ETF survey. [online]. Available at: http://www.ey.com/Publication/vwLUAssets/ey-global-etf-survey-2017/$FILE/ey-global-etf-survey-2017.pdf. [Accessed 8.2.2018]

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, Vol.25 (2), pp. 383-417.

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

Forbes. (2017). The Three largest players have a 70% share in $4 trillion global ETF industry. [online]. Available at: https://www.forbes.com/sites/greatspeculations/2017/05/17/the-three-largest-players-have-a-70-market-share-in-4-trillion-global-etf-industry/#4bc7d75561f6 [Accessed 9.1.2018].

Fulkerson, J. A, Jordan, S. D. and Riley, T. B. (2014). Predictability in Bond ETF Returns. The Journal of Fixed Income, 23(3), pp. 50-63,3.

Gastineau, G. L. (2001). Exchange-Traded Funds: An Introduction. The Journal of Portfolio Management, Vol.27, pp. 88-96.

Goetzmann, W. N., Ivkovic, Z. and Rouwenhorts, K. (2001). Day Trading International Mutual Funds: Evidence and Policy Solutions. Journal of Financial and Quantitative Analysis, 36(3), pp. 287-309.

Hilliard, J. (2014). Premiums and discounts in ETFs: An analysis of the arbitrage mechanism in domestic and international funds. Global Finance Journal Vol.25, pp. 90-107.

Israelsen, C. L. (2005). A refinement to the Sharpe ratio and information ratio. Journal of Asset Management 5, 423–427.

Jares, T. and Lavin, A. (2004). Japan and Hong Kong Exchange-Traded Funds (ETFs): Discounts, Returns, and Trading Strategies. Journal of Financial Services Research, 25(1), pp. 57-69.

Jensen, M. C. (1968). The Performance of mutual funds in the period 1945–1964. Journal of Finance, 23(2), pp. 389-416.

Kreis, Y., Licht, J. and Useche, J. A. (2016). (In)Efficiencies in Latin American ETFs. Cuadernos de Administración, 29(53), pp. 7-48.

Kreis, Y. and Licht, J. (2018). Trading on ETF mispricings. Managerial Finance, 44(3), pp. 357-373.

Levy, A. and Liebermann, O. (2013). Overreaction of country ETFs to US market returns: Intraday vs. daily horizons and the role of synchronized trading. Journal of Banking and Finance, 37(5), pp. 1412-1421.

London Stock Exchange. (2014). ETFs for private investors. [online]. Available at: https://www.lseg.com/sites/default/files/content/documents/ETFs%20for%20Private%20Investors.pdf [Accessed 6.2.2018]

London Stock Exchange. (2017). Exchange traded funds. [online]. Available at: https://www.lseg.com/areas-expertise/our-markets/london-stock-exchange/fixed-income-markets/listed-products/etfs. [Accessed 7.2.2018]

London Stock Exchange. (2018). Prices help. [online]. Available at: https://www.londonstockexchange.com/prices-and-markets/prices-help/prices-help.htm [Accessed 30.12.2018]

Marshall, B. R., Nguyen, N. and Visaltanachoti, N. (2013). ETF arbitrage: Intraday evidence. Journal of Banking and Finance, 37(9), pp. 3486-3498.

Mohamad, A., Jafaar, A. and Goddard, J. (2015). Short selling and exchange-traded funds returns: Evidence from the London Stock Exchange. Applied Economics, pp. 1-13.

Morningstar. (2017) ETF Quickrank. [online]. Available at: http://tools.morningstar.co.uk/uk/etfquickrank/default.aspx?Site=UK&Universe=ETALL%24%24ALL&LanguageId=en-GB&Category=$FOCA$LV$$ [Accessed 10.1.2018].

Newey, W. and West, K. (1987). Hypothesis Testing with Efficient Methods of Moments Estimation. International Economic Review 28, 777–787.

Petäjistö, A. (2017). Inefficiencies in the Pricing of Exchange-Traded Funds, Financial Analysts Journal, vol. 73, issue 1, pp. 24-54.

Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Review of Financial Studies, 22(1), pp. 435-480.

Piccotti, L. R. (2018). ETF Premiums and Liquidity Segmentation. Financial Review, 53(1), pp. 117-152.

Purohit, H. and Malhotra, N. (2015). Pricing Efficiency and Performance of Exchange Traded Funds in India. IUP Journal of Applied Finance, 21(3), pp. 16-35.

Rompotis, G. G. (2010). Does premium impact Exchange-Traded Funds' returns? Evidence from iShares. Journal of Asset Management, 11(4), p. 298.

Ross, S. A., Westerfield, R W. and Jaffe, J. (2013). Corporate finance. 10th ed. McGraw-Hill.

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

Sharpe, W. F. (1966). Mutual Fund Performance. The Journal of Business, 39(1), pp. 119-138.

Simon, D. and Sternberg, J. (2005). Overreaction and Trading Strategies in European iShares. The Journal of Alternative Investments, 8(1), pp. 29-41.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. Journal of Financial Economics, 99(1), pp. 1-10.

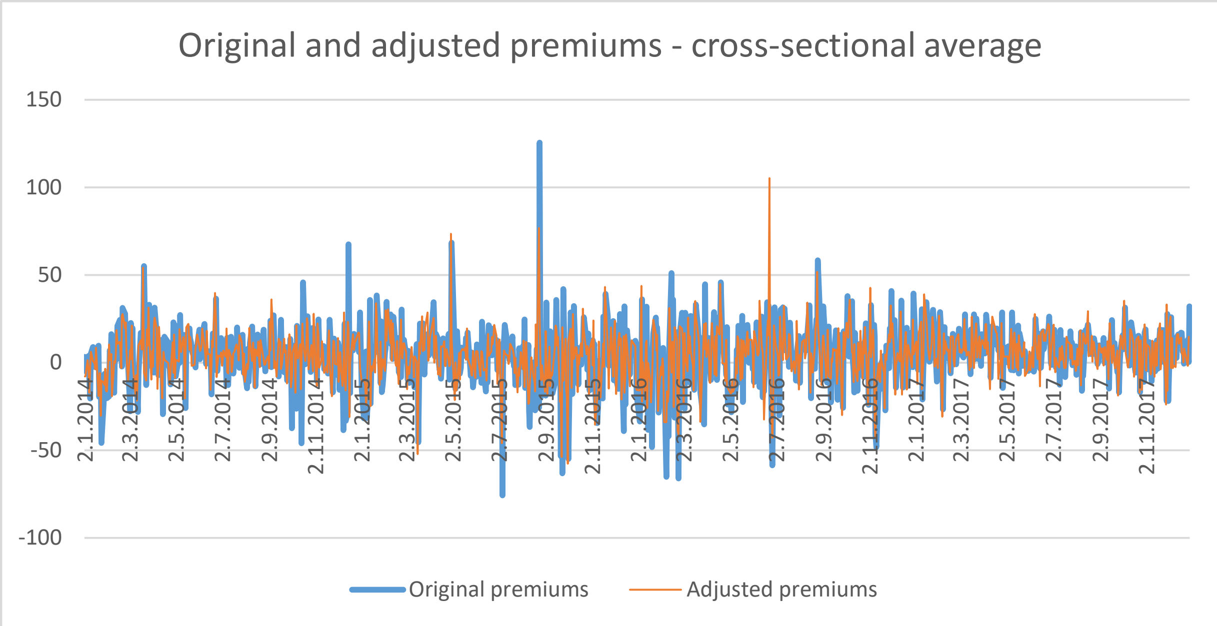
Tripathi, V. and Garg, S. (2016). A Cross-Country Analysis of Pricing Efficiency of Exchange Traded Funds. IUP Journal of Applied Finance, 22(3), pp. 41-63.

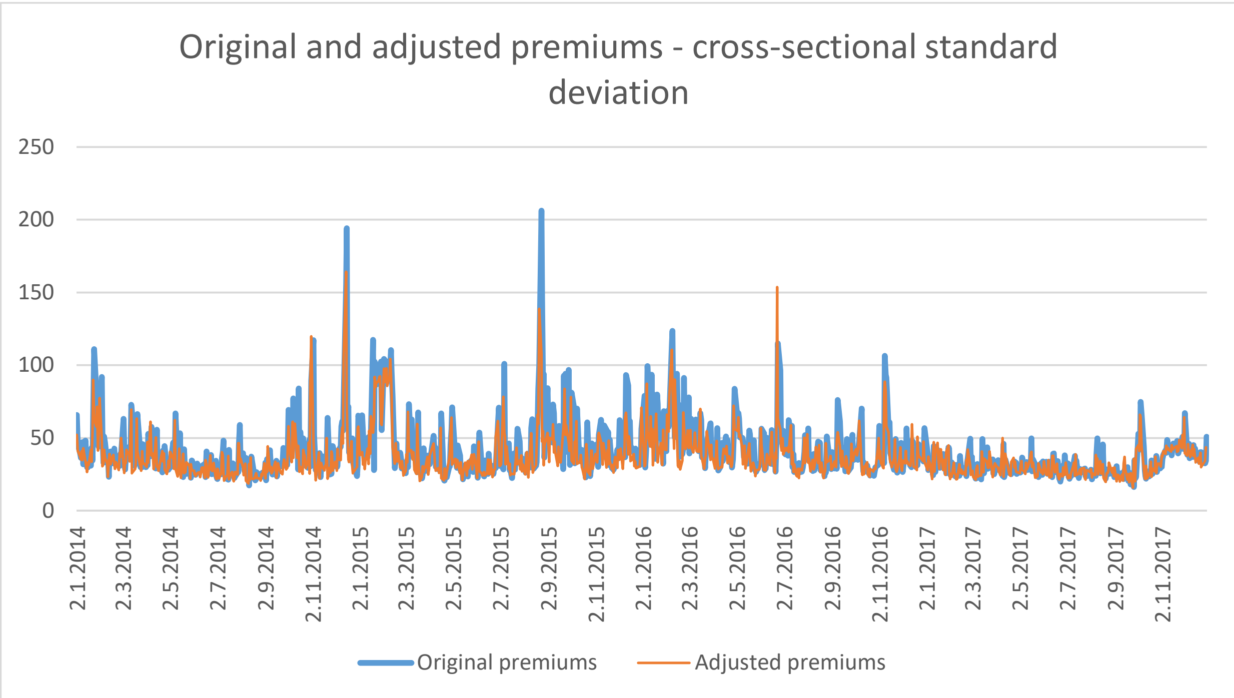
Tse, Y. and Martinez, V. (2007). Price discovery and informational efficiency of international iShares funds. Global Finance Journal, 18(1), pp. 1-15.

Vanguard. (2013). Understanding synthetic ETFs. [online]. Available at: https://pressroom.vanguard.com/content/nonindexed/6.14.2013_Understanding_Synthetic_ETFs.pdf. [Accessed 7.2.2018]

APPENDICES

APPENDIX 1: Cross-sectional comparison of unadjusted premiums and the Engle-adjusted premiums





APPENDIX 2: Pairwise correlations

Pairwise correlation of ETF level factors. Pairwise correlations are calculated to detect multicollinearity. Obtained correlations indicate no multicollinearity.

Variable	Absolute premium	Bid-Ask spread	Volatility of NAV returns	Log of market value	Age in years
Absolute premium	1.000				
Bid-Ask spread	0.167	1.000			
Volatility of NAV returns	0.251	0.082	1.000		
Log of market value	-0.090	-0.274	-0.084	1.000	
Age in years	-0.071	-0.224	0.102	0.500	1.000

Pairwise correlation of market factors. Pairwise correlations are calculated to detect multicollinearity. Obtained correlations indicate no multicollinearity.

Variable	Cross-sectional premium dispersion	FTSE 100 VIX	FTSE All Shares – returns	TED spread
Premium dispersion	1.000			
FTSE 100 VIX	0.505	1.000		
FTSE All Shares – returns	-0.176	-0.128	1.000	
TED spread	-0.109	-0.333	0.054	1.000

APPENDIX 3a: Sub-period analysis of premiums

Sub-period analysis with unadjusted premiums. Table contains statistics for premiums calculated for each ETF investment style. N is the number of observations included within ETF category. Significance of average premium is tested with one sample t-test where hypothesized mean premium is 0. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively.

	Period 2014-2015					Period 2016-2017				
Category	N	Premium avg. (bps)	Volatility (bps)	Min. (bps)	Max. (bps)	N	Premium avg. (bps)	Volatility (bps)	Min. (bps)	Max. (bps)
All	73031	3.27***	52.03	-682.26	1753.85	97892	5.67***	43.96	-819.99	881.04
Equity	51743	2.12***	59.05	-682.26	1753.85	69396	4.22***	48.96	-676.47	881.04
EU	10456	8.43***	24.19	-193.07	665.68	13503	5.26***	21.87	-399.86	881.04
Japan	3036	4.91***	92.68	-592.18	639.74	3449	-0.41	79.17	-485.07	486.32
UK	6461	19.27***	25.89	-682.26	230.87	8389	21.17***	26.28	-197.06	638.40
US	7146	4.62***	84.73	-546.53	1753.85	12969	-1.32***	40.22	-254.12	402.80
Asia	7142	-14.65***	71.56	-451.43	453.03	7932	1.85**	71.85	-676.47	700.53
Emerging	4327	-8.82***	56.20	-309.08	722.77	5308	13.75***	57.80	-497.31	650.95
Latin America	2024	-11.95***	64.19	-333.85	338.66	2020	-28.28***	83.96	-500.95	367.15
Misc. regions	3051	-4.46***	56.78	-411.38	1154.57	3535	3.38***	45.26	-425.68	319.37
World	5186	5.83***	33.68	-154.43	434.07	6254	2.79***	35.83	-385.45	293.41
Sector	2914	-0.15	66.20	-620.07	527.14	6037	0.55	54.11	-641.82	609.44
Bond	15722	6.66***	22.35	-219.36	244.00	22109	11.45***	25.51	-819.99	208.94
Government	7753	-1.67***	20.52	-219.36	244.00	9362	-0.36**	15.68	-134.29	119.56
Corporate general	4896	11.97***	19.88	-144.66	155.27	8495	18.94***	27.28	-819.99	208.94
High-Yield	1555	34.07***	20.57	-142.59	130.90	2737	32.83***	30.11	-219.32	172.01
Diversified	1518	3.95***	9.57	-72.74	63.83	1515	3.72***	10.34	-82.69	108.29
Miscellaneous	5566	4.41***	40.58	-322.39	666.32	6387	1.21***	34.07	-244.49	241.24
Commodity	1012	0.16	46.61	-226.81	275.31	1842	1.90**	36.56	-147.81	241.24
Levered	1518	1.34**	29.34	-148.91	666.32	1515	0.37	23.85	-149.42	152.20
Property	3036	7.35***	42.96	-322.39	309.46	3030	1.21**	36.75	-244.59	206.01

APPENDIX 3b: Sub-period analysis of premiums

Sub-period analysis with the Engle-adjusted premiums. Table contains statistics for the Engle-adjusted premiums calculated for each ETF investment style. N is the number of observations included within ETF category. Significance of average premium is tested with one sample t-test where hypothesized mean premium is 0. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively.

	Period 2014-2015					Period 2016-2017				
Category	N	Engle premium avg. (bps)	Volatility (bps)	Min. (bps)	Max. (bps)	N	Engle premium avg. (bps)	Volatility (bps)	Min. (bps)	Max. (bps)
All	72860	3.13***	44.66	-708.09	1404.59	97860	5.43***	40.63	-662.83	691.02
Equity	51620	1.95***	50.25	-708.09	1404.59	69374	3.93***	44.97	-662.83	691.02
EU	10431	8.37***	23.55	-170.54	607.03	13501	5.40***	21.09	-393.67	658.56
Japan	3030	4.63***	73.63	-512.25	638.32	3448	-2.90**	66.22	-415.58	526.02
UK	6446	19.07***	25.83	-708.09	221.61	8387	20.90***	26.03	-203.88	569.57
US	7125	6.23***	69.75	-508.52	1404.59	12961	-0.05	36.55	-223.68	434.39
Asia	7127	-15.08***	58.57	-378.35	427.98	7931	-0.42	65.24	-662.83	645.33
Emerging	4318	-9.12***	48.61	-309.33	594.23	5305	11.96***	53.56	-469.07	691.02
Latin America	2020	-13.68***	58.69	-292.00	330.27	2020	-28.39***	81.83	-500.77	347.95
Misc. regions	3044	-4.92***	52.61	-395.97	944.14	3535	2.39***	42.35	-421.54	273.24
World	5175	6.16***	29.29	-148.09	459.99	6252	2.93	34.64	-377.8	279.77
Sector	2904	-0.96	55.50	-502.39	444.02	6034	0.99	48.86	-587.54	483.48
Bond	15685	6.61***	22.26	-218.65	233.49	22101	11.30***	25.00	-608.22	203.62
Government	7737	-1.70***	20.36	-218.65	233.49	9359	-0.33**	15.44	-119.90	118.85
Corporate general	4882	11.90***	19.89	-145.87	153.62	8492	18.73***	26.64	-608.22	203.62
High-Yield	1551	33.96***	20.52	-143.94	119.96	2735	32.21***	29.51	-209.24	152.91
Diversified	1515	3.98***	9.51	-70.45	63.58	1515	3.75***	10.21	-81.89	107.84
Miscellaneous	5555	4.34***	35.41	-182.21	646.69	6385	1.34***	31.63	-258.40	293.70
Commodity	1010	-1.60	39.11	-166.44	219.94	1840	2.75***	34.48	-136.99	238.09
Levered	1515	1.44**	28.85	-151.41	646.69	1515	0.50	23.68	-151.62	152.93
Property	3030	7.77***	36.68	-182.21	225.39	3030	0.90*	33.24	-258.40	293.70

APPENDIX 4: Analysis with absolute premiums

Table contains statistics for absolute premiums calculated for each ETF investment style. Both unadjusted and the Engle-adjusted premiums are reported. N of ETFs is the number of funds, while N of obs. is the number of observations included within ETF category. Significance of average absolute premium is tested with one sample t-test where hypothesized mean premium is 0. T-statistics are reported in parentheses. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017.

Category	N of ETFs	N of obs.	absPremium avg. (bps)	Volatility (bps)	Engle N of obs.	Engle absPremium avg. (bps)	Engle Volatility (bps)
All	201	170923	28.62*** (309.00)	38.30	170720	26.31*** (323.88)	33.56
Equity	143	121139	32.59*** (266.37)	42.58	120994	29.57*** (277.67)	37.05
EU	27	23959	14.76*** (121.45)	18.81	23932	14.80*** (127.90)	17.90
Japan	7	6485	58.80*** (75.48)	62.62	6478	47.52*** (74.65)	51.23
UK	17	14850	24.73*** (136.72)	22.01	14833	24.51*** (136.65)	21.84
US	28	20115	32.69*** (92.11)	50.33	20086	28.36*** (94.95)	42.33
Asia	16	15074	49.63*** (115.50)	52.75	15058	43.48*** (116.91)	45.63
Emerging	11	9635	41.99*** (101.94)	40.43	9623	38.17*** (103.82)	36.06
Latin America	4	4044	55.24*** (64.13)	54.78	4040	52.41*** (62.76)	53.08
Misc. regions	7	6586	33.19*** (69.42)	38.81	6579	31.13*** (70.33)	35.90
World	13	11440	22.98*** (92.07)	26.66	11427	21.53*** (93.70)	24.56
Sector	13	8951	36.84*** (77.53)	44.90	8938	33.41*** (81.63)	38.69
Bond	45	37831	16.80*** (163.28)	20.01	37786	16.69*** (165.19)	19.63
Government	19	17115	10.32*** (90.73)	14.86	17096	10.23*** (91.27)	14.66
Corporate general	17	13391	21.28*** (116.84)	21.07	13374	21.25*** (120.17)	20.45
High-Yield	6	4292	35.98*** (100.98)	23.34	4286	35.43*** (100.60)	23.06
Diversified	3	3033	6.55*** (42.73)	8.43	3030	6.46*** (42.38)	8.40
Miscellaneous	13	11953	25.91*** (104.87)	26.96	11940	23.66*** (108.48)	23.83
Commodity	4	2854	29.09*** (55.19)	28.12	2850	26.34*** (56.40)	24.92
Levered	3	3033	17.36*** (46.88)	20.36	3030	17.18*** (47.15)	20.05
Property	6	6066	28.68*** (78.61)	28.36	6060	25.65*** (81.65)	24.45

APPENDIX 5: Persistence analysis in panel data-framework

Premium persistence in panel data-framework. Table contains summary of regressions with model: $PREM_{it} = \alpha + \beta_1 PREM_{it-1} + \beta_2 PREM_{it-2} + \beta_3 PREM_{it-3} + \beta_4 PREM_{it-4} + \beta_5 PREM_{it-} + \varepsilon_{it}$. T-statistics in parentheses are based on the heteroscedasticity robust standard errors. Selection of proper model is derived from the results of the Hausman test followed by the F-test for no fixed effects or the Breusch-Pagan test for random effects. The results of latter two tests are omitted due to space consideration. Categories Latin America, Diversified bonds, Commodity and Levered cannot be tested confidently with the Hausman test or estimated with RE as the total number of ETFs is too low for model with 5 lags. Thus, results of the Hausman test regarding these categories are omitted and model selection is done based on the F-test for no fixed effects. FE-models assume both cross-sectional and period fixed effects. The results are based on the full sample period of 2014-2017. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively.

Category	N	α	β_1	β_2	β_3	β_4	β_5	Adj. R2	Model	Hausman p-value
Equity	120414	1.608*** (11.77)	0.225*** (17.65)	0.109*** (9.68)	0.109*** (11.60)	0.036*** (4.38)	0.048*** (6.00)	0.312	FE	<0.0001
EU	23824	2.791*** (9.54)	0.235*** (5.08)	0.144*** (4.88)	0.101*** (4.40)	0.060*** (3.04)	0.040** (1.96)	0.425	FE	<0.0001
Japan	6450	1.676 (1.56)	0.170*** (7.48)	0.043** (2.47)	0.109*** (6.10)	-0.060*** (-3.70)	-0.024* (-1.69)	0.048	Pooled OLS	0.9380
UK	14765	3.254*** (10.51)	0.257*** (5.68)	0.188*** (4.79)	0.145*** (4.12)	0.133*** (4.64)	0.116*** (4.65)	0.652	FE	<0.0001
US	19970	0.233 (1.12)	0.369*** (3.77)	0.176* (1.90)	0.089 (1.33)	0.064 (1.29)	0.082* (1.88)	0.710	FE	<0.0001
Asia	14994	-1.868*** (-5.26)	0.368*** (19.76)	0.151*** (9.48)	0.062*** (4.11)	0.045*** (3.26)	0.043*** (3.42)	0.728	FE	<0.0001
Emerging	9575	1.081*** (3.72)	0.285*** (10.80)	0.123*** (4.09)	0.134*** (6.16)	0.086*** (4.19)	0.120*** (6.65)	0.753	FE	<0.0001
Latin America	4024	-4.829*** (-5.60)	0.212*** (7.60)	0.196*** (7.08)	0.149*** (5.62)	0.114*** (4.35)	0.097*** (3.81)	0.687	FE	-
Misc. regions	6551	-0.089 (-0.18)	0.154*** (8.06)	0.026 (0.70)	0.042** (2.51)	0.015 (1.01)	0.058*** (3.65)	0.374	FE	<0.0001
World	11375	0.907*** (4.33)	0.375*** (12.98)	0.204*** (7.36)	0.138*** (6.02)	0.048 (1.54)	0.034 (1.19)	0.801	FE	<0.0001
Sector	8886	0.192 (0.36)	0.007 (0.24)	-0.045** (-2.03)	0.052** (2.21)	-0.025 (-1.10)	0.000 (0.02)	0.259	FE	<0.0001
Bonds	37606	1.903*** (15.03)	0.300*** (10.37)	0.177*** (7.81)	0.082*** (6.21)	0.103*** (6.79)	0.053*** (4.17)	0.607	FE	<0.0001
Government	17020	-0.249*** (-2.54)	0.301*** (15.74)	0.196*** (12.54)	0.086*** (5.62)	0.101*** (6.39)	0.070*** (4.45)	0.523	FE	<0.0001
Corporate general	13306	3.932*** (9.41)	0.262*** (4.29)	0.200*** (3.45)	0.110*** (3.29)	0.120*** (3.19)	0.065** (2.19)	0.681	FE	<0.0001
High-Yield	4262	7.051*** (8.51)	0.321*** (11.69)	0.176*** (6.38)	0.102*** (4.07)	0.132*** (5.85)	0.056** (2.46)	0.732	FE	<0.0001
Diversified	3018	3.249*** (11.32)	0.039 (1.27)	0.111*** (3.74)	-0.045 (-1.36)	0.059** (2.11)	-0.013 (-0.47)	0.117	FE	-
Miscellaneous	11888	1.864*** (5.79)	0.107*** (6.72)	0.040*** (2.92)	0.084*** (6.21)	0.035*** (2.77)	0.033** (2.50)	0.194	FE	<0.0001
Commodity	2834	0.920** (2.26)	0.233*** (5.54)	-0.109*** (-3.59)	0.096*** (3.43)	-0.037 (-1.22)	0.037 (1.25)	0.722	FE	-
Levered	3018	0.745 (1.57)	0.069*** (2.75)	0.042** (2.09)	0.027 (1.44)	0.057** (2.16)	-0.034* (-1.76)	0.011	Pooled OLS	-
Property	6036	2.425*** (5.22)	0.163*** (6.82)	0.095*** (4.51)	0.116** (5.31)	0.020 (1.12)	0.035* (1.75)	0.352	FE	<0.0001
All	169908	2.242*** (17.94)	0.216*** (18.92)	0.107*** (10.75)	0.111*** (13.24)	0.041*** (5.54)	0.046*** (6.36)	0.295	FE	<0.0001

APPENDIX 6: Trading strategy with reduced sample

Comparison of trading portfolios. Filter-strategy pursues excess return by taking positions against mispriced ETFs. Filter-strategies are tested with different filters to provide estimates of their profit potential, and in addition, to reveal their sensitivity to transaction costs. Lagged strategies follow one-day waiting rule after signal. Buy-Hold consists of equal investment to each ETF in the given sample. FTSE All shares measures return of investment to local market index. The results are based on the full sample period of 2014-2017. Omitted ETFs (exchange symbols) are BRIC, HMEF, IBZL, LTAM, SPGP.

Sample C: All ETFs (reduced to 196 ETFs)	Filter 125-200	Filter 125-200 (t-costs)	Filter 60-200	Filter 60-200 (t-costs)	Filter 150 lagged	Filter 150 lagged (t-costs)	Buy-Hold	FTSE All shares
Total cumulative return	172.99%	20.86%	417.57%	19.77%	36.95%	-28.12%	34.03%	30.74%
Average daily return	0.104%	0.023%	0.166%	0.021%	0.035%	-0.029%	0.031%	0.030%
Daily standard deviation	0.909%	0.898%	0.771%	0.769%	0.896%	0.893%	0.641%	0.835%
Max	5.457%	5.246%	3.263%	3.109%	9.013%	8.795%	2.650%	3.413%
Min	-5.811%	-6.000%	-3.825%	-3.969%	-6.624%	-6.811%	-3.561%	-4.634%
Sharpe ratio	1.943	0.315	4.111	0.349	0.549	-0.011	0.710	0.495
Filter (+/- bps)	125-200	125-200	60-200	60-200	150	150	-	-
Active trading days	407	407	975	975	322	322	-	-
Transaction cost per roundtrip	-	0.2%	-	0.15%	-	0.2%	-	-

Daily changes in ETF price and NAV. ETFs are assigned based on trading signals on day t. Filter-strategies without transaction costs reported in Appendix 6 are followed. Long position calculates average changes in trading price and NAV between t and t+1 for the ETFs having signal for long positions. Short position calculates the same for the ETFs having signal for short positions. Reported long and short returns are unweighted in contrast to their combined total weighted contributions regarding Filter-strategies. This is seen also in different number of days for each position. One sample t-test, where hypothesized mean is 0, is employed to test the significance of average changes. ***, ** and * represent significances at the 1%, 5% and 10% levels, respectively. The results are based on the full sample period of 2014-2017. Omitted ETFs (exchange symbols) are BRIC, HMEF, IBZL, LTAM, SPGP.

Sample C: All ETFs (reduced to 196 ETFs with filter 125-200)	Long position	Short position
Return ETF	0.0029***	-0.0026**
Return NAV	-0.0083***	0.0101***
N of days	222	227
Sample C: All ETFs (reduced to 196 ETFs with filter 60-200)		
Return ETF	0.0022***	-0.0018***
Return NAV	-0.0041***	0.0035***
N of days	677	875