

**ESSAYS ON EXCHANGE-TRADED FUND MISPRICING AND  
LIQUIDITY**

MARKUS SEBASTIAN BROMAN

A DISSERTATION SUBMITTED TO  
THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BUSINESS ADMINISTRATION  
YORK UNIVERSITY  
TORONTO, ONTARIO

JUNE, 2015

© Markus Broman, 2015

## ABSTRACT

In the first essay I investigate whether the high liquidity of ETFs attracts a clientele of short-term investors. I find that liquidity is an important determinant of fund flows, particularly at weekly and monthly horizons. I also investigate whether more liquid ETFs facilitates shorter-term trading. I document a liquidity clientele amongst institutional investors: i) their buys and sells are positively related to ETF liquidity, ii) liquidity is significantly more important for short-term than for long-term investors, and iii) liquidity is inversely related to average holding periods. These findings are consistent with the idea that liquidity benefits short-term traders the most.

In the second essay I show that changes in ETF misvaluation – as proxied by the return difference between an Exchange-Traded Fund and its underlying portfolio Net Asset Value – comove excessively across ETFs. Excess comovements are positive and highly significant across ETFs in matching investment styles, and negative and significant across distant styles. Further tests based on return reversals suggest that ETF premiums relative to NAV reflect misvaluation primarily in the ETF, rather than the NAV price, particularly for ETFs in more liquid styles (e.g. small-cap). Finally, the degree of return comovements is stronger for funds with high commonality in demand shocks and more attractive liquidity characteristics. These findings are consistent with the idea that liquidity can sometimes be detrimental to pricing efficiency, because liquidity attracts short-horizon noise traders with correlated demand for investment styles.

The third essay investigates whether asset prices are exposed to local-trading induced misvaluation. My sample contains 2727 pairs of European ETFs from 15 different country pairs. Each ETF-pair has identical fundamentals, but is traded by investors from different countries. I document a strong country-specific factor in twin return differentials. This result holds from daily to monthly horizons, it is unrelated to exchange rates and it cannot be explained by liquidity differences. Consistent with local misvaluation in the stock market, I find a significantly positive relationship between twin return differentials and local stock returns, particularly at higher frequencies. Additional tests provide evidence against the alternative explanation that misvaluation is driven by similarities in information diffusion between twins.

## **DEDICATION**

This dissertation is dedicated to my wife Nicole for her unconditional love and support, and to my daughter Emily for constantly reminding me what is important in life.

## **ACKNOWLEDGEMENTS**

The second chapter in this dissertation is the recipient of Eastern Finance Association (EFA) 2014 Outstanding Doctoral student award. The third chapter is a semifinalist for the Best Paper Award in Investments at the Financial Management Association (FMA) annual meeting. Thank you for recognizing my research. Many thanks to Pauline Shum for her encouragement, support and advice, to Yisong Tian for our stimulating discussions and to Kee-Hong Bae for his invaluable feedback on my research.

Special thanks to Larry Harris and Francesco Franzoni for their insightful comments and suggestions on my first essay, and Jaewoon Choi for his feedback on my second essay. Others to whom I am grateful are Mark Kamstra, Sandy Suardi, Anna Agapova, Jared DeLisle, Dennis Bams, Kristian Miltersen, Thierry Foucault, Lee Cohen, Daniel Andrei and Akiko Watanabe.

I am also very grateful for comments received at York University (Canada), Ryerson University, Norwegian School of Economics, Aalto University School of Business, ESSEC Business school, Syracuse University, the 2013 European Financial Management Association meeting (EFMA), the 2014 Eastern Finance Association, the 2014 European Financial Management Association (FMA Europe), the 2014 European Finance Association conference (EFA), the 2014 Northern Finance Association conference (NFA), the 2014 FMA Annual Meeting (FMA) and the 2015 American Finance Association conference (AFA). Finally, I am indebted to faculty, staff and students at York for their help and support.

## TABLE OF CONTENTS

<b>ABSTRACT .....</b>	<b>ii</b>
<b>DEDICATION .....</b>	<b>iii</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>iv</b>
<b>CHAPTER ONE: INTRODUCTION AND MOTIVATION .....</b>	<b>1</b>
1.1 INTRODUCTION.....	1
<b>CHAPTER TWO: SHORT-TERM TRADING AND LIQUIDITY CLIENTELES: EVIDENCE FROM EXCHANGE-TRADED FUNDS .....</b>	<b>8</b>
2.1 INTRODUCTION.....	8
2.2 DATA.....	12
2.2.1 Defining liquidity.....	14
2.2.2 Descriptive statistics .....	18
2.3 ETF DEMAND AND LIQUIDITY.....	19
2.3.1 Share creations and ETF flows.....	19
2.3.2 Descriptive statistics: Liquidity, share creations and ETF flows.....	20
2.3.3 Determinants of net fund flows.....	21
2.3.4 Baseline results: ETF flows .....	24
2.3.5 Results: Liquidity and ETF flows .....	24
2.3.6 Institutional demand.....	26
2.3.7 Results: Institutional demand and liquidity .....	28
2.4 LIQUIDITY CLIENTELES AND SHORT-TERM TRADING .....	29
2.4.1 Institutional ownership – baseline model.....	29
2.4.2 Determinants of short- vs. long-term institutional ownership .....	30
2.4.3 Short-term trading and liquidity .....	33
2.5 SUMMARY AND CONCLUSION .....	35
<b>CHAPTER THREE: LIQUIDITY, STYLE INVESTING AND EXCESS COMOVEMENT OF EXCHANGE-TRADED FUND RETURNS .....</b>	<b>50</b>
3.1 INTRODUCTION.....	50
3.2 BACKGROUND ON ETF ARBITRAGE AND INSTITUTIONAL DETAILS .....	55
3.2.1 Limits-to-arbitrage .....	56
3.3 THEORETICAL FRAMEWORK AND TESTABLE IMPLICATIONS.....	57
3.3.1 Key variables .....	61
3.4 DATA.....	61
3.4.1 Descriptive statistics .....	63
3.5 EMPIRICAL TESTS OF EXCESS COMOVEMENT.....	64
3.5.1 Results: style-based commonality in misvaluation.....	66
3.5.2 Results: sub-samples .....	67
3.5.3 Robustness: Exposure to systematic risk.....	68

3.6 SOURCE OF MISVALUATION: ETF OR NAV?.....	70
3.7 CORRELATED DEMAND AND EXCESS COMOVEMENT IN RETURNS.....	72
3.7.1 Measuring correlated demand shocks .....	72
3.7.2 Results: commonality in demand shocks .....	75
3.7.3 Explaining the amount of commonality in misvaluation.....	76
3.8 SUMMARY AND CONCLUSIONS.....	79
<b>CHAPTER FOUR: LOCAL TRADING HABITATS AND COMMONALITY IN MISVALUATION: EVIDENCE FROM TWIN-SECURITIES.....</b>	<b>90</b>
4.1 INTRODUCTION.....	90
4.2 LOCAL TRADING HABITATS AND TESTABLE HYPOTHESES.....	95
4.2.1 Identifying the local habitat factor and testable implications .....	96
4.2.2 Key variables .....	99
4.3 DATA.....	101
4.3.1 Descriptive statistics .....	102
4.4 EMPIRICAL TESTS OF LOCAL HABITATS .....	103
4.4.1 Results: composite local habitat factor.....	104
4.4.2 Separating the local habitat factors for $L$ and $F$ .....	106
4.4.3 Placebo test .....	107
4.5 LOCAL MISVALUATION IN THE STOCK MARKET.....	108
4.5.1 Constructing local stock market factors .....	109
4.5.2 Results: misvaluation and local stock market returns.....	110
4.5.3 Robustness: regional risk factors.....	111
4.6 FURTHER EVIDENCE OF MISVALUATION VS. INFORMATION DIFFUSION .....	111
4.6.1 Liquidity and commonality in misvaluation.....	113
4.6.2 No information advantage.....	115
4.6.3 Return predictability and the reversal of misvaluation.....	116
4.7 SUMMARY AND CONCLUSIONS .....	118
<b>CHAPTER FIVE: CONCLUSION .....</b>	<b>130</b>
<b>BIBLIOGRAPHY.....</b>	<b>137</b>
<b>APPENDICES.....</b>	<b>144</b>
APPENDIX A3-1: VARIABLE DESCRIPTIONS .....	144
APPENDIX A4-1: ETF-DATA CLEANING.....	145
APPENDIX A4-2: STOCK-DATA CLEANING.....	146
APPENDIX A4-3: THE MECHANICS OF ARBITRAGE AND CROSS-LISTINGS WITH EUROPEAN ETF .....	147

## LIST OF TABLES

Table 2- 1: Descriptive statistics .....	39
Table 2- 2: Correlations among liquidity variables.....	40
Table 2- 3: ETF share creation activity.....	41
Table 2- 4: Measures of ETF demand and liquidity .....	42
Table 2- 5: Explaining ETF inflows – base model .....	43
Table 2- 6: Inflows and liquidity.....	44
Table 2- 7: Horizon of net flows and liquidity .....	45
Table 2- 8: Institutional trading and liquidity.....	46
Table 2- 9: Institutional ownership and ETF Liquidity.....	47
Table 2- 10: Determinants of short- vs. long-term institutional ownership .....	48
Table 2- 11: Short-term trading and liquidity .....	49
Table 3- 1: Snapshot of ETF statistics.....	82
Table 3- 2: Descriptive Statistics.....	83
Table 3- 3: Style-based commonality in misvaluation.....	84
Table 3- 4: Sub-sample results for commonality in misvaluation.....	85
Table 3- 5: Differences in systematic risk? .....	86
Table 3- 6: Source of mispricing: ETF or NAV?.....	87
Table 3- 7: Comovement in relative trading activity and liquidity.....	88
Table 3- 8: Explaining the degree of excess return comovement .....	89
Table 4- 1: Selecting ETF twins .....	120
Table 4- 2: Snapshot of ETF sample 11/2013 .....	121
Table 4- 3: Summary statistics on ETF misvaluation .....	122
Table 4- 4: Trading location specific misvaluation .....	123
Table 4- 5: Separating the local and foreign habitat factors .....	124
Table 4- 6: Placebo test for local misvaluation in ETF returns.....	125
Table 4- 7: Twin return differentials and local stock market returns.....	126
Table 4- 8: Degree of commonality in twin return differentials and fund-pair liquidity.....	127
Table 4- 9: Does local comovement reflect local information?.....	128
Table 4- 10: Return predictability and reversal of mispricing.....	129

## LIST OF FIGURES

<b>Figure 1- 1: Global assets invested in ETFs/ETPs and hedge funds, at the end of Q1 2015 .....</b>	<b>7</b>
<b>Figure 2- 1: Time-series variability of relative liquidity (Balanced Panel) .....</b>	<b>37</b>
<b>Figure 2- 2: Time-series Variability of Relative Institutional Ownership.....</b>	<b>38</b>
<b>Figure 3- 1: Commonality in misvaluation, commonality in demand shocks and the level of liquidity .....</b>	<b>81</b>

## **CHAPTER ONE: INTRODUCTION AND MOTIVATION**

### **1.1 INTRODUCTION**

As of August 2014, the global Exchange-Traded Fund (ETF) market has garnered \$2.7 trillion of assets, similar in size to the global hedge fund industry, despite the fact that hedge funds predate ETFs by 41 years. Figure 1-1 illustrates the growth in assets of the ETF and the hedge fund industry, side-by-side from 1993 to 2014. The ETF industry reached \$600 billion as late as in 2006, roughly six years later than the hedge fund industry. Although the U.S. leads in assets with \$430 billion in 2006 and \$2 trillion in 2014, European ETF assets have also grown rapidly from \$92 billion in 2006 to \$500 billion in 2014<sup>1</sup>. In terms of trading volume, ETFs already account for 42 percent of all equity dollar volume on NYSE ARCA<sup>2</sup>.

The advantages of ETFs are well known. They offer investors tax-efficient exposure to a diversified basket of securities at real-time prices, and at a lower fee than their retail index fund counterparts. In addition, In Chapter two, I show that many ETFs are more liquid relative to their

---

<sup>1</sup> Source: The data on the ETF vs. hedge-fund comparison is from ETFGI Research Note: “ETF/ETP industry forecast to surpass hedge funds in Q2”, April 2015. Data on U.S. ETF assets is from the Investment company fact book 2005. The data on European AUM is from Financial times, “Chart that tells a story — ETF growth”, May 1, 2015.

<sup>2</sup> “Measuring ETF Liquidity: Looking beyond Trading Volume” SPDR University, SPDR.com

underlying portfolio of securities. For instance, between 2006 and 2012 the SPDR S&P 500 ETF (ticker: SPY) the proportional quoted spread of SPY was only 18 percent of those for its underlying basket, and its Amihud's price impact was only 17 percent of those for its basket. This is far from an isolated incident. For our sample of U.S. equity ETFs (consisting of 165 ETFs with \$540 billion in assets, or 85 percent of U.S. equity ETF assets), relative liquidity in spreads is on average 89 percent of the underlying portfolio spreads. Moreover, the ETF structure provides an extra (or "hidden") layer of liquidity that allows select institutional investors (known as Authorized Participants, or APs) to create or redeem ETF shares for their retail or institutional clients via the primary market whenever it is more cost efficient, i.e. when the liquidity of ETF in the secondary market is below that for its underlying securities. This option-like feature of ETFs implies that the lower bound for their liquidity is determined by the maximum of the liquidity of the ETF in the secondary market, and the liquidity of the underlying securities in the primary market.

In the *second chapter*, I examine whether improved liquidity of ETFs relative to its underlying basket (if any) leads to greater net demand for ETFs, and encourages shorter-term trading. Our conjecture is that liquidity facilitates short-term trading, in which case liquidity should be an important determinant of ETF net demand over shorter horizons. Liquidity should be important not only for buying, but also selling (Cella, Ellul and Giannetti (2013), for example, discuss the importance of liquidity particularly for institutional selling, especially during episodes of market distress, such as the financial crisis of 2008.). I measure ETF demand either by net fund flows (over weekly, monthly or quarterly horizons), or institutional demand (net, buy or sell) constructed from quarterly institutional holdings from SEC 13-F filings.

Although liquidity is beneficial to all investors, short-term investors should value liquidity the most (Amihud and Mendelson, 1986). Recent survey evidence by Greenwich Associates (2013) shows that institutional investors frequently use ETFs for short-term trading strategies: as a short-term hedging vehicle, for transition management, for tactical asset allocations, and cash equitization. Amongst institutional investors, I therefore study further whether there is a liquidity clientele effect, whereby more liquid ETFs (relative to their underlying basket) experience increased ownership by shorter-term investors relative to longer-term, and reduced holding periods.

Short-term trading and the existence of liquidity clienteles is a critical assumption in the third chapter, where I argue that ETFs attract a clientele of short-term investors with correlated non-fundamental demand for investment styles. More generally, there is also an ongoing debate on whether short-term trading is beneficial to, or detrimental for, market efficiency. The empirical evidence is mixed, possibly because short-term trading can theoretically be beneficial when it is related to arbitrage activities (Shleifer and Summers, 1990, for example); detrimental when it reflects common non-fundamental demand such as style switching (Barberis and Shleifer, 2003), sentiment (DeLong, Shleifer, Summers, and Waldmann, 1990), or correlated liquidity needs (Greenwood and Thesmar, 2011) that generate excess comovements among asset returns; or detrimental when it reflects exogenous trading needs such as income shocks (Lynch and Tan, 2007) or liquidity shocks (Huang, 2003) that increase the frequency of trading and impose a liquidity risk factor on asset prices.

In the *third chapter* I investigate how style investing and the high liquidity of ETFs can facilitate excess comovements in ETF returns. In frictionless markets with rational investors, the price of a security will equal its fundamental value, and any comovement in returns must be due to comovement in fundamentals. However, in economies with frictions or with irrational investors, and in which there are limits to arbitrage, comovement in returns may be partially delinked from fundamentals giving rise to what is known as *excess comovement*. Several theoretical models build on the price impact of correlated investor demand to explain the existence of excess comovement (Barberis and Shleifer, 2003; Barberis, Shleifer and Wurgler, 2005; and Greenwood and Thesmar, 2011).

The theoretical channel for excess comovement in ETF returns relies the model by Barberis and Shleifer (2003), where investors allocate funds at the style level (e.g. small or value) as opposed to at the individual asset level, moving into styles that have performed well in the past, and out of styles that have performed poorly. If some of these style investors are also noise traders with correlated sentiment (e.g. Baker and Wurgler, 2006), then coordinated shifts in investor preferences across investment styles (e.g. from value to growth) will induce a common factor in the returns of assets in the same style.

Investor demand should go first to the securities where the purest play exists and where liquidity is highest. Exchange-Traded Funds provide investors with easy access to popular investment styles (large, small, value, growth, sector) at a cost that is on average lower relative to

their underlying basket of securities (as shown in Chapter 2). Moreover, it is easy to move money in and out of two different styles with ETFs and to enter into long-short strategies (e.g. Value-Growth) due to the relatively low short-selling costs of ETFs. My conjecture is that, due to the ease of investing in investment styles with ETFs and because of their high liquidity, ETFs attract a clientele of short-term investors with correlated non-fundamental demand for investment styles. In this case, liquidity can sometimes be detrimental to pricing efficiency, because the high liquidity of ETFs is attractive to short-term investors, which increases the exposure of ETFs to common non-fundamental demand shocks.

To arrive at a testable hypothesis, I make a comparison between the return of the ETF and the return of its underlying portfolio Net Asset Value (NAV). Since both are claims on the same underlying cash-flows, the ETF-NAV return difference can be used as a proxy for the change in ETF misvaluation. This relative, or twin-based, comparison allows me to identify excess comovements by studying common factors in the change in misvaluation. This approach is in sharp contrast to existing studies that investigate anomalous return comovements around “exogenous” events, or by relying on a CAPM type model to filter out the fundamental component of returns. Moreover, by properly controlling for fundamental drivers of return comovements, I can examine what affects the degree of excess comovements in order to provide a better understanding of the ETF characteristics, particularly liquidity, that drive a wedge in the clientele between ETFs and their underlying securities.

Common factors in ETF misvaluation (ETF-NAV returns) can also arise for reasons other than correlated demand shocks. Specifically, if fundamental news about abstract risk-factors is incorporated first into ETF prices (the high liquidity of ETFs is attractive to long-term investors as well), while their underlying securities react more slowly, then similarities in information diffusion across ETFs can generate excess comovements in ETF returns (Barberis, Shleifer and Wurgler, 2005). In order to disentangle the two hypotheses, I use a direct approach by examining the source of misvaluation (ETF vs. NAV). The non-fundamentals based view predicts that ETFs are hit by temporary demand shocks that subsequently revert, in which case current ETF premiums (ETF-NAV price deviations) should predict future ETF returns negatively, while future NAV returns remain unaffected. In contrast, the information diffusion view predicts that ETFs impound fundamental information first, while their underlying securities (NAV) catch up

with a lag, in which case current ETF premiums (ETF-NAV price deviations) predict future NAV returns positively, while future ETF returns are not affected.

In the *fourth chapter* I analyze whether excess comovements can arise in an international setting, specifically, whether the trading by a local (country-specific) clientele in Exchange-Traded Funds can induce a country-specific factor in ETF misvaluation. My focus is on ETFs tracking European equity indices and with a European stock market listing. A unique advantage of the European ETF market (in terms of research design) is the fragmentation of listings across various exchanges. Although there are more than 650 plain vanilla European equity ETFs traded in Europe, many are essentially duplicates of one another as they are either cross-listed from a different European exchange, or they track an identical benchmark index. For instance, the most popular benchmark index, Euro Stoxx 50, is tracked by 78 different ETFs. ETF providers list new (or cross-list existing) funds primarily to meet local clientele demand. Out of the 416 ETFs and six countries analyzed in this chapter (U.K., Switzerland, France, Germany, Italy and Netherlands), I construct 2727 (local-foreign) pairs of ETFs from 15 different country pairs. Each ETF in a pair has identical fundamentals (i.e. passively tracks an identical benchmark index), but is traded and owned by investors from different countries.

Locally excessive return comovements can arise theoretically in the preferred habitat view of excess comovement formalized by Barberis, Shleifer and Wurgler (2005), where a) some investors limit their trading to a particular subset of securities, such as local stocks or ETFs; or b) some securities are primarily owned by a specific subset of investors, such as local investor. The correlated trading of these investors can then generate a common factor in the returns of securities in the habitat. A similar home bias in “foreign investments” has previously been documented by Ke, Ng and Wang (2010), who find that non-US-based mutual funds prefer to invest in locally traded securities of U.S. companies.

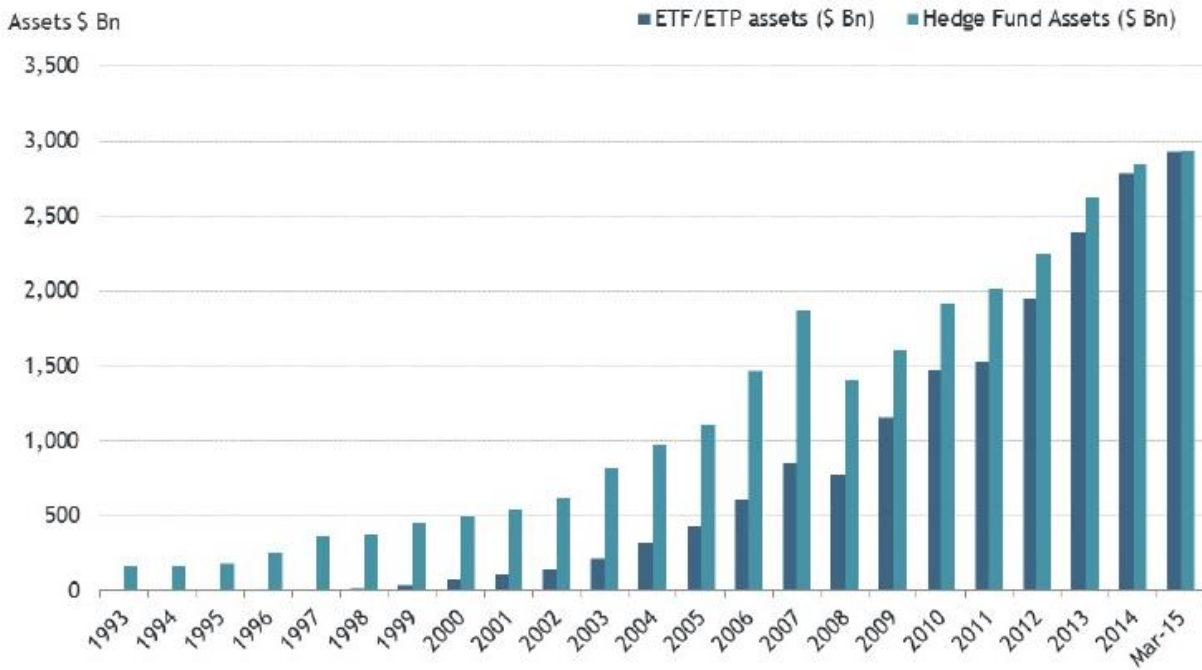
The motivation for local clienteles in ETFs is closely related to the literature on home and local bias. Home bias, or the tendency for investor to invest in equity securities at home, is pervasive around the world (e.g. Chan, Covrig an Ng, 2005), while local bias reflects the tendency for investors to invest in nearby domestic securities (e.g. Coval and Moskowitz, 2001). Local ETF habitats can arise theoretically in the framework of Chan, Covrig an Ng (2005). They show theoretically that home bias can arise if local investors have lower deadweight costs when investing in the domestic market (or higher costs when investing abroad) relative to the world

investor. Locally traded ETFs provide investors with easy access to foreign stock market indices without having to incur costs associated with the trading, settlement or transfer of ETF shares across national markets, operational costs that arise from the requirement to have depository accounts in each country and follow post-trading rules in each jurisdiction. Other costs include withholding taxes, which apply to dividend income that foreign investors (non-residents) pay to the government where the stocks are domiciled, and possibly also to the government where the ETF is located. International trading restrictions and lack of information (Merton, 1987) or constraints on investors' attention (Peng and Xiong, 2006) can also contribute to the formation of local habitats.

I also investigate an alternative explanation of excess comovements that arises from differences in the speed of information diffusion between securities traded in different countries (Barberis, Shleifer and Wurgler, 2005). Specifically, if local investors are better or faster at processing information about abstract local fundamental factors, then it can give rise to excess comovements among local securities relative to their foreign peers. To rule out this alternative explanation, I compare the strength of excess comovements among sub-samples of ETFs where local investors are unlikely to have an informational advantage over foreign investors. As a more direct test, I also examine the source of mean-reversion in ETF misvaluation. Local-trading induced misvaluation suggests that current ETF premiums predict future returns of locally traded ETFs negatively, while the information diffusion story predicts that current ETF premiums predict future returns of foreign ETFs positively, as foreign ETF prices catch up to reflect the fundamental news already incorporated into foreign ETF prices.

*The fifth chapter* provides a summary of key results and conclusions.

**Figure 1- 1: Global assets invested in ETFs/ETPs and hedge funds, at the end of Q1 2015**



**Source:** ETFGI research note 27 April 2015, <http://etfgi.com/news/industry/newsid/162>

## **CHAPTER TWO: SHORT-TERM TRADING AND LIQUIDITY CLIENTELES: EVIDENCE FROM EXCHANGE-TRADED FUNDS**

### **2.1 INTRODUCTION**

*“Unlike traditional index funds, ETFs trade on an exchange, and can be bought and sold at any time. That creates a temptation to trade...”*

~ John Bogle, Founder of the Vanguard Group<sup>3</sup>

As of August 2014, the global Exchange-Traded Fund (ETF) market has garnered \$2.7 trillion of assets,<sup>4</sup> similar in size to the global hedge fund industry,<sup>5</sup> despite the fact that hedge funds predate ETFs by 41 years. Their trading volume accounts for 42 percent of all equity dollar volume on the NYSE Arca by about the same time.<sup>6</sup> ETFs offer investors tax-efficient exposure to a diversified basket of securities at real-time prices, and at a lower fee than their retail index fund counterparts. In addition, many ETFs are more liquid than their underlying basket of securities; that is, they have positive *relative* liquidity. One dramatic example in terms of the

---

<sup>3</sup> Newlands, Chris. “Bogle launches renewed attacks on ETFs” Financial Times, March 15, 2015.

<sup>4</sup> Including both ETFs and ETPs, from just \$426 billion in 2005. Source: ETFGI database.

<sup>5</sup> The amount of global hedge fund assets under management was \$2.82 trillion in the third quarter of 2014. Source: Hedge Fund Research Global Hedge Fund Report Third Quarter 2014. It is widely expected that ETFs will eclipse hedge funds by the third quarter of 2015.

<sup>6</sup> “Measuring ETF Liquidity: Looking beyond Trading Volume” SPDR University, SPDR.com.

proportional bid-ask spread is the SPDR S&P 500 ETF (ticker: SPY). Between 2006 and 2012, SPY was trading at spreads of approximately 18 percent of those for its underlying basket. Other examples include iShares Russell 2000 (ticker: IWM), the SPDR S&P Midcap 400 ETF (ticker: MDY), the iShares Russell 2000 Growth (ticker: IWO) and Value (ticker: IWN), all of which were trading at spreads of 11 to 42 percent of those for their respective underlying baskets. In terms of trading volume, the ETF structure provides an extra (or “hidden”) layer of liquidity, allowing institutional investors to create or redeem ETF shares via the primary market whenever it is more cost efficient, or when there is a demand and supply imbalance.

In this paper, I examine whether improved liquidity of ETFs relative to its underlying basket (if any) leads to greater demand for ETFs, and encourages shorter-term trading. I measure liquidity in four ways. The first three relate to secondary market liquidity: quoted spreads, price impact, and turnover. The fourth relates to the creation and redemption process in the primary market. The decisions that I am considering are that of institutional investors, who have the ability to choose between investing in an ETF and its underlying basket. For most retail investors, the transaction costs involved in trading the underlying basket are likely prohibitive. Therefore, relative liquidity may be less important for retail investors. Our conjecture is that liquidity facilitates short-term trading, in which case liquidity should be an important determinant of ETF demand over shorter horizons. I measure ETF demand either by net fund flows (over weekly, monthly or quarterly horizons), or institutional demand (net, buy or sell) constructed from quarterly institutional holdings from SEC 13F filings. Amongst institutional investors, I study further whether there is a liquidity clientele effect, whereby more liquid ETFs (relative to their underlying basket) experience increased ownership by shorter-term investors relative to longer-term, and reduced holding periods.

There is an ongoing debate on whether short-term trading is beneficial to, or detrimental for, market efficiency. The empirical evidence is mixed,<sup>7</sup> possibly because short-term trading can theoretically be *beneficial* when it is related to arbitrage activities (Shleifer and Summers, 1990, for example); *detrimental* when it reflects common non-fundamental demand such as style switching (Barberis and Shleifer, 2003), sentiment (DeLong, Shleifer, Summers, and Waldmann, 1990), or correlated liquidity needs (Greenwood and Thesmar, 2011) that generate excess

---

<sup>7</sup> Several papers have linked short-term trading to more anomalous pricing of stock returns, while others have linked short-term institutional trading to greater pricing efficiency (see Cremers and Pareek (2014) for an extensive list of references).

comovements among asset returns; or detrimental when it reflects exogenous trading needs such as income shocks (Lynch and Tan, 2007) or liquidity shocks (Huang, 2003) that increase the frequency of trading and impose a liquidity risk factor on asset prices.

Although liquidity is beneficial to all investors, short-term investors should value liquidity the most (Amihud and Mendelson, 1986). Recent survey evidence by Greenwich Associates (2013) shows that institutional investors frequently use ETFs for short-term trading strategies: as a short-term hedging vehicle, for transition management, for tactical asset allocations, and cash equitization. The incidence of short-term trading in ETFs is also evident from a comparison of holding durations of ETFs relative to their underlying securities. To illustrate, using quarterly holdings data from SEC 13F filings, I find that the mean holding period for our sample of ETFs is 2.8 quarters, which is on average two quarters *lower* than the dollar-weighted holding duration of its underlying basket. (Because I do not have access to daily or intraday holdings data, I am not looking at arbitrage transactions or high frequency trading of ETFs.)

In order to provide a clean comparison of liquidity across ETFs and underlying securities, I focus on a select sample of 165 ETFs that are traded in the U.S. and that track only U.S. equity indices. In other words, both the ETFs and their underlying securities are traded in the same time zone, and in the same currency. The total assets under management of our sample as of the end of 2012 is \$540 billion - roughly 85 percent of all U.S. equity ETFs. Within this universe, investors can choose from a wide range of funds tracking size, valuation and sector-based indices.

If liquidity facilitates short-term trading, then fund flows at higher frequencies should be more responsive to liquidity than at lower frequencies. I quantify ETF, underlying basket, and relative liquidity (ETF minus underlying portfolio liquidity) using the four measures mentioned previously. To preview our results, I show that in general, ETF (underlying basket) liquidity predicts future fund flows positively (negatively), or alternatively, that relative liquidity predicts net flows positively. These results hold for weekly and monthly flows, while for quarterly flows liquidity is no longer a significant determinant. In contrast, trading costs that are more relevant for longer-term investors – namely expense ratios – are more than 10 times stronger for quarterly net flows as opposed to weekly. The relationship between liquidity and future fund flows also remains highly significant when I use unexpected liquidity shocks instead of liquidity levels.

Net ETF flows capture aggregate demand, both by retail and institutional investors. To investigate the hypothesis that relative liquidity matters primarily for institutional investors, I begin by constructing a measure of net institutional demand, and separate measures for institutional buying and selling. Net institutional demand is the change in ETF shares held by 13-F investors over a quarter relative to shares outstanding at the beginning of the quarter, while institutional buying (selling) is the aggregate number of shares bought (sold) by 13-F investors during a quarter. Similarly to quarterly net flows, I do not find a significant relationship between liquidity and net institutional demand. Although it is not surprising to find that liquidity matters less over longer horizons, a calendar quarter is still a relatively short holding period and I should not expect the significance to disappear altogether. Indeed, our results demonstrate that ETF liquidity is a strong predictor of institutional buying, and an even stronger predictor institutional selling (Cella, Ellul and Giannetti (2013), for example, discuss the importance of liquidity particularly for institutional selling, especially during episodes of market distress, such as the financial crisis of 2008), which is consistent with the idea that short-term trading strategies require liquidity both for buys and sells.

In terms of more direct evidence of a liquidity clientele effect amongst institutional investors, I show that liquidity is more important for investors with short-term trading needs. First, institutional ownership is significantly higher among ETFs with higher relative liquidity. Indeed, institutional investors account for more than 80 percent of the trading in ETFs (Aggarwal and Schofield, 2012), while in our sample their total ownership is on average roughly 40 percent, which suggests that institutions are more active traders in ETFs. Second, I compare the fraction of short-term ownership in ETFs, using the definition of transient investors by Brian Bushee (1998), with the fraction of short-term ownership in the underlying securities, and find that ETFs have on average 13.1 percent more short-term investors. Third, I look at institutional ownership by short- and long-term investors and show that relative liquidity is two to six times as important for short- than for long-term institutional ownership (significant at the one percent level).

Further, in order to show the incidence of short-term trading in ETFs, I use Cremers and Pareek's (2014) measure of institutional investors' average quarterly holding duration (which I call ETF duration). This measure is unrelated to arbitrage activities because it is calculated from quarterly portfolio holdings. It therefore ignores arbitrage trades in response to ETF mispricing, as the latter is short-lived. I also investigate the determinants of ETF duration and find that ETF

liquidity is amongst the most important explanatory factor. These results confirm that ETFs are used for short-term trading, which is facilitated by the high liquidity of ETFs. Therefore, an important implication of our findings is that as the ETF market continues to grow at a rapid pace, it will likely facilitate a higher volume of short-term trading, thereby affecting market efficiency and volatility.

The reverse causality story – i.e., greater use of ETFs by short-term investors leads to higher ETF liquidity – is theoretically also a possibility. If true, it could explain the result that liquidity predicts more short-term trading. The empirical evidence suggests, however, that greater ownership by short-term investors can sometimes lead to lower (rather than higher) liquidity. For example, Cella, Ellul and Giannetti (2013) show that during episodes of market turmoil, liquidity dry-ups are exacerbated for stocks held primarily by short-term investors. In this case the empirical finding that liquidity encourages shorter-term trading is inconsistent with reverse causality.

This study proceeds as follows. Section 2 describes the data, defines the key variables and presents the summary statistics. The empirical analysis in section 3 focuses on the relationship between liquidity and ETF demand, measured either by net fund flows (over weekly, monthly or quarterly horizons) or by quarterly institutional demand (net, buys or sells) using 13-F data. Section 4 provides more direct evidence on liquidity clienteles by investigating the importance of liquidity for institutional ownership, and for short- vs. long-term investors, and whether average institutional holding periods are directly related to ETF liquidity. Section 5 concludes.

## **2.2 DATA**

Our data selection starts with all U.S. traded ETFs. I keep funds that i) invest only in U.S. equity, ii) are physically replicated and passively managed, and iii) have at least 3 years of data available.<sup>8</sup> I exclude actively managed ETFs because investors presumably choose them for their portfolio strategy and manager performance, not necessarily for reasons of liquidity and short-term trading. Further, the segment of active funds is still very small, representing about one percent of total ETF assets. In recent years, a number of “smart beta” funds have emerged. I also exclude these funds because they are semi-active strategies. These three exclusion criteria

---

<sup>8</sup> We exclude the first 6 months of a fund's history given that liquidity measures based on shares outstanding are prone to extreme outliers as newly created ETFs experience dramatic creation/redemption activity early in its lifecycle.

decrease the sample of ETFs from 363 to 224 to 165. In terms of assets under management (AUM), the total AUM of U.S. equity funds was \$632 billion in September 2012 according Blackrock (2012), while the AUM of our 165 funds was \$540 billion, or roughly 85 percent of the AUM of all U.S. equity ETFs. Given the dramatic expansion in the scope and size of the ETF market in the last few years, earlier data on ETF creations is likely to suffer from non-stationarity. Only around 2006-2007 has the growth in AUM started to stabilize and creations have become more stationary. This is why I decided to focus on a more recent sample period, from January 2006 to December 2012.<sup>9</sup>

Our sample of ETFs, NAV prices, underlying indices and number of shares outstanding<sup>10</sup> are obtained from Bloomberg. As reported by Petajisto (2011), Bloomberg coverage of ETFs is anywhere from 60 to 97 percent of all ETFs by number, and 90 to 99 percent by assets. ETF price, volume and return data are obtained from CRSP.

Next, I use Morningstar Direct to access several fund characteristics and ETF holdings, which are used for computing measures of an ETF's underlying basket's liquidity. The holdings data in Morningstar is either monthly or quarterly (the latter in the case of Vanguard ETFs, for example). In order to identify a benchmark index's investment style, I use the ubiquitous Morningstar style classification system that characterizes funds along two dimensions: small-, mid- and large-cap as well as value, growth and blend. I do this for three reasons. First, the Morningstar classification coincides with dichotomy often used by practitioners. Second, Morningstar is a leading fund information provider and its classification system is publicly available. Third, many ETFs are named after their Morningstar style analogs e.g. SPDR S&P 600 Small-Cap Value or iShares Russell 3000 Growth fund. The style classifications are updated monthly. Finally, I also use yearly net expense ratios from Morningstar.

At the end of 2012, our raw sample consists of 165 U.S. equity ETFs; 80 core (or non-sector) and 85 sector ETFs. Among the core styles, there is roughly an equal distribution of value-growth-blend funds within each size-category, except the small-value (with 3 ETFs) and large-blend (21 ETFs) categories (see Table 2-1). The most popular sector styles are financial, technology and real-estate with 13, 12 and 11 ETFs respectively. At the start of our sample period in January 2006, the total AUM of the sample is \$211.27 billion, the mean (median)

---

<sup>9</sup> Prior to 2000, there were very few ETFs, most of them were single foreign country index funds.

<sup>10</sup> The number of shares outstanding in Bloomberg is updated daily whereas CRSP generally updates shares outstanding once a month. Therefore, many share creations/redemptions are not captured by CRSP.

AUM is \$2.2 billion (\$0.46 billion). The total AUM grows to \$544 billion at the end of our sample period in December 2012, and the mean (median) is \$3.17 billion (\$0.59 billion). As the means and medians indicate, the AUM is highly skewed in the cross-section of funds.

I obtain institutional ownership data for our sample of ETFs from the CDA/Spectrum database provided by Thomson Reuters. This data comes from quarterly 13-F filings with the U.S. Securities and Exchange Commission (SEC). After 1978, all institutional investment managers (including foreign investors) with discretionary assets in excess of \$100 million in Section 13(f) securities (mainly publicly traded equity) are required to report their holdings to the SEC on a quarterly or semi-annual basis. SEC regulation stipulates that all holdings of common stock (including ETFs) greater than 10,000 shares or \$200,000 must be disclosed.

Following Frazzini and Lamont (2008), the holdings data is adjusted for stock splits, and I discard observations where the report date is subsequent to the file date, or when the number of shares in a fund's portfolio exceeds the total amount of shares outstanding. The data has also been cleaned for inconsistencies and missing data has been filled in using prior and future holdings observations as outlined in Campbell, Ramadorai and Schwartz (2009).

Finally, I also use disaggregated institutional ownership by type of institution from Thomson Reuters. The type codes from this data set have two critical problems. First, as noted on the WRDS website, the type code variable is not reliable after 1998 due to an internal mapping error. I address this issue by using the classification codes provided by Brian Bushee. Second, the manager number (*MGRNO*) is not always unique to a manager/institution. Instead, I use the permanent keys provided by Bushee<sup>11</sup>.

### **2.2.1 Defining liquidity**

Liquidity is a multi-faceted concept and capturing it with a single measure is likely inadequate. I therefore focus on four dimensions of liquidity: Amihud's price impact, quoted spreads, turnover, and share creation/redemption activity. First, Amihud's price impact is an important measure particularly for institutional investors whose trades are typically large and which are known to affect stock prices (see e.g. Edelen and Warner, 2001). I define the monthly price impact (*AMIHUD*) for an ETF or a constituent stock as:

---

<sup>11</sup> We thank Brian Bushee for making his data available to us at: <http://acct.wharton.upenn.edu/faculty/bushee/IIclass.html>.

$$AMIHUD_{i,t} = \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \frac{|R_{i,d}|}{DVOL_{i,d}} \right] \quad (1)$$

where  $R_{i,d}$  = closing mid-quote return (in %) for ETF or stock  $i$  on trading day  $d$ ;  
 $DVOL_{i,d}$  = dollar volume (in \$ millions) for ETF or stock  $i$  on trading day  $d$ ;  
 $N_t$  = number of trading days in calendar month  $t$ .

I take the natural logarithm of the daily price impact before calculating the monthly average in order to reduce the impact of outliers in the data. This is particularly important when Eq. (1) is calculated for stocks (Karolyi, Lee and Van Dijk, 2012). For consistency, I also take the natural logarithm of our other liquidity measures. Amihud's price impact has been widely used in the literature. Hasbrouck (2009) reports that, "among the daily proxies, the Amihud illiquidity measure is most strongly correlated with the TAQ-based price impact coefficient" (p. 1459). Several other papers also endorse Amihud's measure as a good proxy for price impact (e.g. Lesmond, 2005; Goyenko, Holden and Trzcinka, 2009; Fong, Holden, and Trzcinka; 2010).

Second, the proportional quoted spread ( $QSPR$ ) is a direct measure of liquidity for normalized transactions and it features in many theoretical models on liquidity and asset pricing (see e.g., Amihud and Mendelson 1986). The monthly average  $QSPR$  for an ETF or a constituent stock is defined as follows:

$$QSPR_{i,t} = \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \frac{ASK_{i,d} - BID_{i,d}}{(ASK_{i,d} + BID_{i,d})/2} \right] \quad (2)$$

where  $ASK_{i,d}$  and  $BID_{i,d}$  = ask and bid price at the close on trading day  $d$  for ETF or stock  $i$   
 $N_t$  = number of trading days in calendar month  $t$ .

Third, I look at turnover. The definition for turnover ( $TO$ ) for an ETF or a constituent stock is:

$$TO_{i,t} = \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \frac{VOL_{i,d}}{SHR_{i,d}} \right] \quad (3)$$

where  $VOL_{i,d}$  = share volume of ETF or stock  $i$  on day  $d$ ;  
 $SHR_{i,d}$  = number of shares outstanding;  
 $N_t$  = number of trading days in calendar month  $t$ .

Our argument for using turnover as a measure of liquidity is based on the notion that turnover measures the intensity of trading activity, and is therefore related to holding periods. As Amihud

and Mendelson (1986) point out, turnover is inversely related to illiquidity costs.<sup>12</sup> Before computing the monthly average measures, I discard daily price impact and turnover observations beyond the 99.5<sup>th</sup> percentile of the distribution to mitigate the impact of extreme outliers.

Unlike traditional equities, one concern with ETFs is that their turnover and price impact measures contain information about trading volume only from the secondary market. In contrast, shares that are created (or redeemed) in response to buy-and-hold demand may bypass the secondary market and go directly to (from) the investor's portfolio. In the case of turnover, it moves in the *opposite* direction relative to the activity in the primary market. Specifically, a share creation that bypasses the secondary market will result in a decline in turnover as the number of shares outstanding increases, and vice-versa for redemptions. Trading volume in the secondary market is affected only when the creation/redemption arises from arbitrage trades.<sup>13</sup> Even in this case, there may be a discrepancy because arbitrage trades may have been initiated any time within the 3-day settlement period, while new shares show up in the data only at settlement. Given that I average the liquidity measures over a monthly horizon, I do not expect the aforementioned time-series phenomena to affect our results. Nevertheless, I may gain some additional insights by considering a liquidity measure that is specifically designed to capture activities in the primary market, which leads us to our fourth measure of liquidity: share creation/redemption activity (henceforth share creation), defined as:

$$CREATE_{i,t} = \log \left[ 1 + \frac{1}{N_t} \sum_{d=1}^{N_t} \frac{|SHR_{i,d} - SHR_{i,d-1}|}{SHR_{i,d-1}} \right] \quad (4)$$

where:  $SHR_{i,d}$  = shares outstanding for ETF  $i$  on day  $d$ ;  
 $N_t$  = number of trading days in calendar month  $t$ .

A high share creation activity is attractive to investors, particularly institutional investors, because it proxies for primary market liquidity, and captures the additional layer of liquidity that ETFs possess. For example, new ETF shares can be created to meet excess demand and to minimize price impact in the secondary market, if the underlying securities have sufficient liquidity. This option-like feature of ETFs implies that the lower bound for their liquidity is

---

<sup>12</sup> Other explanations for turnover have also been proposed, such as investor disagreement, investor attention and the speed of adjustment to market-wide information (see e.g. Cremers and Pareek, 2014).

<sup>13</sup> For example, in a creation, there is no impact on trading volume if an investor already has the underlying securities in an existing portfolio, and is keeping the newly created ETFs.

determined by the maximum of the liquidity of the ETF in the secondary market, and the liquidity of the underlying securities in the primary market.

Following Lo and Wang (2000), Yan (2008), Idzorek *et al.* (2012), I estimate the liquidity of the *underlying* portfolio by dollar-weighting each security's liquidity (i.e. the weight of the security in the ETF's basket). In our case, I use the three secondary market measures presented above: Amihud's price impact (1), quoted spreads (2), and turnover (3).

The liquidity of an ETF's underlying portfolio plays two roles in determining demand. First a more liquid underlying portfolio will facilitate the creation and redemption process of the ETF, thereby improving the ETF's primary market liquidity. Second, an institutional investor can gain exposure to a stock index by either purchasing the ETF or the constituents directly. In making this decision, the investor may compare the liquidity of the two options. Therefore, I also consider an ETF's relative liquidity, in price impact, quoted spreads and turnover, defined as:

$$\begin{aligned}
REL(AMIHU_{i,t}) &= \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \sum_{k=1}^K w_{i,k,d} \frac{|R_{k,d}^{UND}|}{DVOL_{k,d}^{UND}} / \frac{|R_{i,d}^{ETF}|}{DVOL_{i,d}^{ETF}} \right] \\
REL(QSPR_{i,t}) &= \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \sum_{k=1}^K w_{i,k,d} QSPR_{k,d}^{UND} / QSPR_{i,d}^{ETF} \right] \\
REL(TO_{i,t}) &= \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \sum_{k=1}^K w_{i,k,d} \frac{VOL_{k,d}^{UND}}{SHR_{k,d}^{UND}} / \frac{VOL_{i,d}^{ETF}}{SHR_{i,d}^{ETF}} \right]
\end{aligned} \tag{5}$$

where  $w_{i,k,d}$  = dollar-weight, or the weight that each security has in the underlying basket held by the ETF, of stock  $k$  in ETF  $i$  on day  $d$ .

I construct the relative liquidity measures in a consistent manner across all three measures. First, I take the daily log-difference between ETF and underlying portfolio liquidity. Second I take the monthly average of this measure. Since I am using log-differences, the relative liquidity measures can be interpreted as the percentage difference in liquidity between the ETF and its underlying portfolio. As for share creation activity (*CREATE*), I can already view it as a measure of relative, rather than just ETF liquidity since the underlying stocks have no regular share creation activities.

In the empirical tests I present results either for all liquidity measures, in which case I draw inferences that apply broadly across several measures, or I use a composite measure given that previous research has highlighted that variations in liquidity are governed by a significant

common component.<sup>14</sup> The composite measure is constructed using the first principal component for i) the ETF [Eq. (1) to (4)], ii) the underlying portfolio [Eq. (1) to (3)], and iii) relative liquidity measures [Eq. (4) and (5)].

## 2.2.2 Descriptive statistics

To illustrate the dynamics of relative liquidity, I graph the different measures for a balanced panel of ETFs that have existed since 2006.<sup>15</sup> Figure 2-1 shows that relative spreads are roughly zero for the average fund, but varies considerable over time. If I look across funds grouped by underlying index size (pooling core and sector styles together based on Morningstar's 3-by-3 size and valuation style), I can see that small-cap ETFs, followed by mid- and by large-cap, offer the highest relative liquidity in spreads. After the financial crisis in 2008, the underlying basket spreads of small-cap ETFs have been around 50 to 100 percent higher than that for the ETF itself. As for Amihud's price impact, Figure 2-1 shows an increasing trend over time in the large- and small-cap size categories. The average small-cap fund has consistently provided positive relative liquidity via price impact, while for the average large-cap funds relative liquidity is negative. As shown in Figure 2-1, both relative turnover and share creation activity have generally declined over time for the average fund, except toward the end of 2008, at the peak of the financial crisis. This observation is likely a reflection of the balanced panel of funds maturing: younger funds typically have higher relative turnover and share creation activity relative to older funds (at the median).

[Figure 2-1]

Table 2-1 reports summary statistics for several additional variables of interest. Blend or core funds are on average at least twice as large as value or growth funds, while differences in medians are less pronounced. In terms of expense ratios, the notable difference is between funds with exposure to sector and non-sector funds, with the former's median ratio being more than twice as high. I will discuss the last two columns in Table 2-1 in a later section.

[Table 2-1]

---

<sup>14</sup> See e.g., Chordia, Roll, and Subrahmanyam (2000); Amihud, (2002); Korajczyk and Sadka (2008)

<sup>15</sup> New ETFs typically have low relative liquidity in spreads and price impact, but high relative liquidity in turnover and share creation activity. For this reason, we illustrate relative liquidity using a balanced panel in Figure 1. In the panel regression analysis, we will employ the full sample.

The correlations between the different relative liquidity measures are positive, but far from perfect: the range is from 0.19 to 0.77 (see Table 2-2). *Relative* price impact and spreads are the most highly correlated (0.77), while relative spreads and share creation activity are the least correlated (0.19). ETF liquidity and relative ETF liquidity measures are highly correlated (from 0.65 to 0.95), because most of the variations in relative liquidity (ETF minus underlying portfolio liquidity) are driven by ETF liquidity. In robustness tests, I also use shocks to liquidity. The correlations here are from -0.02 to 0.65.

[Table 2-2]

## 2.3 ETF DEMAND AND LIQUIDITY

In this section, I focus on the importance of liquidity for ETF demand, measured either by net ETF flows (sections 3.1-3.5) or quarterly institutional buys and sells (sections 3.6 and 3.7). Our conjecture is that liquidity is more important for shorter-term demand (weekly and monthly net flows as opposed to quarterly), and for both buys and sells because short-term trading strategies require liquidity in both instances.

### 2.3.1 Share creations and ETF flows

ETF data on share creations are available daily, because they are traded like a stock. Therefore, net ETF flows can be estimated at the daily interval using the change in the number of shares outstanding and the NAV per share on the day of the (net) share creation. That said, I will focus primarily on monthly flows, derived from daily flows, although in some tests I also use weekly or quarterly flows. The reason for aggregating daily flows is that net share creations are lumpy: due to a minimum creation/redemption size, they tend to happen less frequently and in large batches. Also, daily flows are prone to extreme outliers, as evidenced by the difference in the means and the medians: in our sample, the average (median) ETF has net share creations on 30.9 percent (22.7 percent) of the trading days, and conditional on such days, the magnitudes are \$69.6 million (\$12.4 million) or 2.8 percent (1.0 percent) of AUM (see Table 2-3).

I define the net fund flow over the horizon  $t$  (= week, month or quarter) as:

$$IFL_{i,t} = \frac{\sum_{d=1}^{N_t} (SHR_{i,d} - SHR_{i,d-1}) NAV_{i,d}}{AUM_{i,t-1}} = \frac{\sum_{d=1}^{N_t} (AUM_{i,d} - AUM_{i,d-1}(1 + r_{i,d}))}{AUM_{i,t-1}} \quad (6)$$

where  $SHR_{i,d}$  = number of shares outstanding for ETF  $i$  on day  $d$ ;

$$\begin{aligned}
NAV_{i,d} &= \text{Net Asset Value per share;} \\
r_{i,d} &= \text{NAV return over day } d; \\
AUM_{i,t-1} &= \text{Assets Under Management at the end of period } t-1.
\end{aligned}$$

The second expression of inflows in Eq. (6) corresponds to what is typically used in the mutual fund literature with one exception. The dollar flow in the numerator is measured at the daily level and is aggregated over the horizon  $t$  to obtain per-period dollar flows. This definition is more accurate because ETF flows are measured using actual prices ( $NAV_{i,d}$ ) observed on the day when flows occurs. In contrast, the monthly flow typically used in the mutual fund literature (i.e.  $AUM_{i,t} - AUM_{i,t-1}(1 + r_{i,t})$ ), assumes that flows occur at the end of each period (see for example, Sirri and Tufano, 1998). The reason is that mutual fund data are available only at the monthly interval, unlike ETF data which can be obtained daily. In order to mitigate the impact of outliers and potential data errors, I winsorize net fund flows at the 1<sup>st</sup> and 99<sup>th</sup> percentiles.

### 2.3.2 Descriptive statistics: Liquidity, share creations and ETF flows

If liquidity facilitates short-term trading, then fund flows at higher frequencies should be more responsive to liquidity than at lower frequencies. In addition, the magnitude of share creations/redemptions (i.e. ignoring the sign) should be positively related to liquidity because it is beneficial both for buying (creations) and selling (redemptions).

First, I provide summary statistics on the frequency of share creations/redemptions, as well as the dollar and proportional magnitudes. The relationship between daily share creations and the secondary market liquidity measures Eq. (5) can be seen when I condition on the AUM of the funds. To simplify the presentation, I group ETFs into a high and low liquidity category where the cut-off is based on the median liquidity rank at a given point in time.<sup>16</sup> The results in Table 2-3 show that within each AUM tercile, funds with above median relative liquidity have more share creation activity relative to funds below the median. To illustrate, in the middle AUM-tercile, funds with high relative liquidity have, on average, more frequent share creations (on 34.6 percent versus 20.2 percent of days), larger dollar magnitudes (\$20.0 million versus \$9.3 million) as well as greater relative magnitudes (4.1 percent versus 2.4 percent of AUM, or

---

<sup>16</sup> Specifically, each ETF is ranked based on the equally-weighted liquidity rank across the three measures of relative liquidity in Eq. (5) in the prior month. Share creation activity is excluded here because it is mechanically related to the magnitude of share creations.

equivalently of shares outstanding) compared to funds with low relative liquidity. These differences are significant at the 5 percent level both at the mean and the median.

[Table 2-3]

Second, I look at the relationship between net fund flows and relative liquidity measured over different horizons (see Table 2-4). The results suggest that *weekly* fund flows are on average 0.48 percent across all funds (the median is zero due to infrequent share creation), but 0.55 percent for funds with high relative liquidity and 0.40 percent for funds with low relative liquidity. The difference is significant at the 5 percent level. Similarly, at the *monthly* horizon, the median difference between funds with high relative liquidity (0.67 percent) vs. low relative liquidity (0.00 percent) is significant at the 1 percent level. The difference in means is also positive, but insignificant. However, at the quarterly horizon there are no significant differences in fund flows between the two liquidity categories, consistent with the notion that liquidity matter more at higher frequencies where short-term trading is more prevalent. Another distinct feature is that funds with higher liquidity have much more volatile flows.

I also provide a breakdown of net fund flows by relative liquidity (high vs. low) conditioning on fund size (high, medium and low AUM). At the weekly and monthly horizons, net fund flows are significantly higher for more liquid funds relative to less liquid funds, but only among medium and low AUM funds. Amongst the largest funds, liquidity no longer matters.

[Table 2-4]

### 2.3.3 Determinants of net fund flows

I model net fund flows as follows:

$$IFL_{i,t} = a_0 + b_1 LIQ_{i,t-1} + b_2 Fund_{i,t-1} + b_3 Control_{t-1} + FE + \varepsilon_{i,t} \quad (7)$$

where  $LIQ_{i,t-1}$  = ETF & underlying portfolio or relative liquidity during period  $t-1$   
 $Fund_{i,t-1}$  = vector of fund characteristics  
 $Control_{t-1}$  = vector of control variables, mainly macro variables  
 $FE$  = fixed effects: year, month, sector, style and/or fund

i)  $LIQ_{i,t-1}$

I include measures of ETF, underlying portfolio or relative liquidity to investigate whether this critical determinant of short-term demand is important for net ETF flows.

ii)  $Fund_{i,t-1}$

First, I consider ETF mispricing. ETFs can be mispriced at any point in time, trading at a premium to its fundamental value, or at a discount. The level of mispricing is defined as the log deviation of the ETF price from its NAV:

$$PREM_{i,d} = \ln(ETF_{i,d}) - \ln(NAV_{i,d}) \quad (8)$$

where  $ETF_{i,d}$  = ETF<sub>i</sub>'s bid-ask midpoint on day  $d$ ;  $NAV_{i,d}$  = Net Asset Value of ETF<sub>i</sub>

In calculating the premium, I use the bid-ask midpoint as suggested by Engle and Sarkar (2006). I expect ETF flows to be positively associated with premiums because arbitrageurs can take advantage of positive (negative) premiums by creating new (redeeming old) ETF shares. Specifically, as the ETF price rises above its NAV (i.e.,  $PREM > 0$ ), APs can execute an arbitrage trade and profit from the creation of new ETF shares, through an exchange of the underlying basket with the ETF manager, and vice versa when  $PREM < 0$ . The extent to which premiums forecast flows over longer horizons is, however, unclear. Much of the arbitrage activity, APs and investors such as hedge funds engaging in long-short ETF strategies, occurs daily or intra-daily. In unreported results, I find that daily premiums only persist for a few days, but average premiums over weekly or monthly horizons contain a strong persistent component. Hence, I use the average daily premium over the period  $t$  ( $PREM_{i,t}$ ) to capture excess demand for an ETF that has yet to be arbitrated away.

Second, Barber, Odean and Zheng (2005) show that salient, attention-grabbing information such as front-end loads and commissions, are important for mutual fund investors' purchase decisions, while less transparent operating fees are not important. However, Grinblatt *et al.* (2013) provide direct evidence that high-IQ investors prefer mutual funds with lower fees relative to low-IQ investors using a comprehensive dataset of Finnish males. For ETFs, the direct costs include commissions and expense ratios. Commissions for ETFs have gone down dramatically; several trading platforms now offer ETF trading free of commissions.<sup>17</sup> With this in mind, I only control for a fund's annual net expense ratio in the fund flow regressions.

Third, tracking error, measured by the standard deviation of the daily change in premiums,  $PREM_{i,d}$ , (measured over the same horizon as for ETF flows) may also be an

---

<sup>17</sup>A list of commission-free ETFs can be found at <http://etfdb.com/type/commission-free/all/>. Free commissions often come with restrictions, such as limits on how quickly a transaction can be turned over.

important determinant of ETF flows when investors are concerned about the magnitude of mispricing, whether positive or negative. Traditionally, tracking error is based on the return differences between the NAV and the underlying index, which reflects the ETF manager's replication strategy and performance, rather than between the ETF price and its NAV (Eq. (8)). Our definition is slightly different, and it reflects the ability of the APs in executing arbitrage trades to eliminate mispricing (i.e., limits to arbitrage).

Fourth, I consider past returns as a predictor of fund flows. Existing theoretical studies on mutual fund flows typically focus on the pursuit of superior talent (see Gruber, 1996; Berk and Green, 2004; Zheng, 1999), in which case investors rationally select managers based on their past performance. In the context of passively managed ETFs, the search for alpha cannot explain return chasing behaviour. Alternatively, return chasing can reflect positive feedback trading in investment styles, which arises in the theoretical style chasing model of Barberis and Shleifer (2003) or in herding models where investors learn from each other's signals (see e.g. Banerjee, 1992), or follow correlated signals (see Froot, Scharfstein, and Stein, 1992; Hirshleifer, Subrahmanyam, and Titman, 1994). In light of the recent evidence that ETF investors engage in positive feedback style chasing (Broman and Shum, 2015b), I control for up to four lags of ETF returns as well as four lags of own and distant style returns based on Morningstar's 3-by-3 size and valuation index.

Fifth, ETF inflows may be affected by a number of other fund-specific characteristics, such as age ( $AGE$ ,  $AGE^2$ ), cumulative inflows and size ( $AUM$ ).

iii)  $Control_{t-1}$

I also include a number of control variables to capture aggregate market conditions: the VIX index, Pastor and Stambaugh (2003) level of market liquidity and the term spread between the 10-year and 3-month U.S. government yields. For additional details on the variable definitions, see Appendix 1.

iv)  $FE$

I estimate regression (7) using pooled OLS with fixed effects for style (Morningstar 3-by-3 categories), sector (0, 1), time (year, month separately) and in some cases, also fund effects. Standard errors are clustered by calendar time to account for cross-sectional dependence in the residual demand shocks.

### 2.3.4 Baseline results: ETF flows

I first present results for the baseline model (Eq. (7) without the liquidity variables). The results in Table 2-5 show that ETF flows (passive funds), unlike mutual fund flows, are not persistent: the first lag is negative and significant while the remaining lags are positive but generally insignificant. Similar results are also obtained for quarterly net flows (unreported to conserve space), and for another measure of ETF demand, namely quarterly net institutional demand calculated from 13F holdings (more details in section 3.6). The lack of persistence in ETF flows is likely a reflection of how ETFs are used and created. For instance, it may reflect the outcome of arbitrage trades, or short-term style chasing strategies (Broman and Shum, 2015b).

Specification (2) in Table 2-5 also shows that the other fund-specific characteristics enter with the expected signs. Bigger and older funds, and funds with higher expense ratios grow more slowly. Flows also increase during periods of high market volatility (VIX) and when the term-spread is high. These findings indicate that passive investing using ETFs grows during uncertain times. This may reflect in part a shift away from active vehicles, and/or the desire to switch to ETFs because they are perceived to be more liquid during uncertain times, which would be particularly attractive for short-term traders. Specifications (3) and (4) provide sub-sample results for non-sector and sector funds respectively. The main difference between the two is that the mean-reversion in flows is mainly evident for sector funds, and macro variables are generally less significant for sector funds.

[Table 2-5]

### 2.3.5 Results: Liquidity and ETF flows

In Table 2-6 I look at the impact of liquidity on fund flows. Because of the high correlations between the liquidity measures (see Table 2-2), I perform separate regressions for each measure. For consistency, I *sign* all of the variables such that larger values indicate higher liquidity. Our first set of results in Table 2-6 suggest that ETF liquidity is positively and significantly related to fund flows (all measures), while underlying portfolio liquidity enters negatively, but significantly only for price impact. The weaker results for underlying portfolio liquidity may reflect a lack of variations in this measure, because some funds share similar underlying portfolios (e.g., a broad U.S. index), and especially since I have controlled for several related macro variables, such as

Pastor and Stambaugh (2003)'s market-wide liquidity level, term-spread and the VIX, as well as year and month dummies.

[Table 2-6]

Given the opposite signs of the ETF and underlying portfolio liquidity variables, the implication is that investors prefer highly liquid ETFs that invest in securities with lower liquidity. Hence, relative liquidity, as a separate variable, ought to be able to explain this result. Indeed, this is what I find. Fund flows are higher for funds with higher *relative* liquidity in spreads, price impact, turnover and share creation activity (all significant at 5 percent level or better). To illustrate the economic magnitude of these effects, I compare the impact of relative liquidity between funds at the 25<sup>th</sup> and the 75<sup>th</sup> percentiles of the distribution. The increase in monthly flows within the interquartile range is 0.51 percent, 2.14 percent, 1.43 percent and 0.98 percent respectively for *REL(QSPR)*, *REL(AMIHU)*, *REL(TO)* and *CREATE*, respectively. These numbers are economically meaningful when compared to the mean (median) monthly flow of about 2.15 (0.00) percent. The explanatory power of the other fund-specific variables is not materially affected by the inclusion of the liquidity variables. In unreported tests I also confirm that similar results are obtained in sub-samples of funds based on age and *AUM*.

The results so far have shown that funds with higher liquidity have higher future flows. This is largely a cross-sectional phenomenon given the high degree of persistence in our liquidity measures. I also investigate the importance of the time-series dimension of liquidity on flows by replacing the liquidity measures with shocks to liquidity (from an AR(1) model). This specification, reported in Panel B, is also less prone to concerns about reverse causality – whereby persistent demand creates liquidity – given that neither flows, nor liquidity shocks are persistent. I also add ETF fixed effects in order to focus on time-series variations in liquidity. The results show that shocks to relative liquidity (via price impact, turnover and share creation activity) predict ETF flows positively and significantly (at the 1 percent level or better). In contrast to the previous results, quoted spreads are no longer significant. The implication is that while both time-series and cross-sectional variations in liquidity matter for predicting ETF flows, the cross-section seems to be more important.

Next, I investigate the importance of liquidity on net fund flows measured over different horizons and for sub-samples of core versus sector funds (Table 2-7). To keep the table compact,

I combine the different liquidity measures into one by using the first principal component (separately for ETF and underlying portfolio). If liquidity is more important for short-term trading, then the impact of liquidity on flows should be stronger at shorter-horizons. Indeed that is what I find. ETF liquidity enters with a positive and significant coefficient at the weekly and monthly horizons, while it is insignificant at the quarterly horizon. The results are similar regardless of whether liquidity levels (Panel A) or liquidity shocks (Panel B) are used. Although the results are highly significant both for core and sector ETFs, the magnitude of the coefficients of ETF liquidity are between three to five times higher for sector funds, which may reflect the popularity of short-term sector rotation strategies among ETF users. Furthermore, sector ETFs have significantly shorter holding periods relative to core ETFs (see Table 2-1, to be discussed in greater detail in section 4).

Another indication that quarterly flows are less likely to reflect short-term trades is the negative and highly significant coefficient on net expense ratios, which are more relevant for longer-term investors. Since I have already controlled for ETF fixed effects, this result implies that an increase in expense ratios is negatively associated with future net fund flows. I also estimate regressions with style fixed effects instead of ETF fixed effects in order to better identify the impact of expense ratios. These tests, which are unreported for conciseness, show that while expense ratios enter with a negative and significant coefficient across all horizons, the magnitude of the effect is roughly 10 times higher in quarterly as opposed to weekly net flows. In this specification ETF (underlying portfolio) liquidity enters with a positive (negative) and (generally) significant coefficient, consistent with relative liquidity driving the results.

[Table 2-7]

Overall, the results in this section provide further evidence that liquidity is an important determinant of flows at higher frequencies, both among core and sector funds, as well as across time and in the cross-section of funds.

### **2.3.6 Institutional demand**

Net ETF flows capture aggregate demand, both by retail and institutional investors. In contrast, our conjecture is that relative liquidity matters primarily for institutional investors, who have the ability to choose between investing in an ETF and its underlying basket. For most retail

investors, the transaction costs involved in trading the underlying securities are likely prohibitive. Therefore, relative liquidity may be less important relative to institutional investors.

In order to specifically investigate the demand by institutional investors, I construct a measure of net institutional demand (*NID*) based on the total number of ETF shares bought or sold by all 13F investors during quarter  $q$  scaled by the number of shares outstanding at  $q-1$ :

$$NID_{i,q} = \frac{\sum_{j=1}^{N_j} (Shares_{i,j,q} - Shares_{i,j,q-1})}{SHR_{i,q-1}} \quad (9)$$

where  $Shares_{i,j,q}$  = number of shares owned by 13F investor  $j$  at the end of quarter  $q$   
 $SHR_{i,q-1}$  = total number of shares outstanding of ETF  $i$  at the end of quarter  $q-1$

*NID* is positive when institutional investors buy more shares than they sell from retail investors (or other non 13F investors), and/or when institutional investors create new ETF shares in excess of what they redeem. Thus, *NID* captures demand both in the primary and secondary markets.<sup>18</sup>

I also calculate institutional demand separately for buys (*IB*) and sells (*IS*):

$$IB_{i,q} = \frac{\sum_{j=1}^{N_j} (Shares_{i,j,q} - Shares_{i,j,q-1}) I[Shares_{i,j,q} - Shares_{i,j,q-1} > 0]}{SHR_{i,q-1}} \quad (10)$$

$$IS_{i,q} = \frac{\sum_{j=1}^{N_j} -(Shares_{i,j,q} - Shares_{i,j,q-1}) I[Shares_{i,j,q} - Shares_{i,j,q-1} < 0]}{SHR_{i,q-1}} \quad (11)$$

where  $I[.]$  = an indicator variable for buys (8) or sells (9) by institution  $j$

The prediction is that liquidity is an important predictor of both institutional buying (*IB*) and selling (*IS*) since short-term trading requires liquidity in both instances, particularly because institutional investors have an incentive to exit during episodes of market turmoil when overall liquidity is already depressed (Cella, Ellul and Giannetti, 2013). It is possible some of the liquidity proxies may also be correlated with arbitrage activity (e.g. higher turnover or share creation activity), which would predict both more institutional buying and selling (but no impact on net demand) because mispricing is symmetric and on average zero, but varies in the short-run. This is, however, not a concern here because mispricing is very short-lived: a regression of the daily ETF premium (8) on its lags shows that only the first few daily lags are statistically

---

<sup>18</sup> Note that *NID* is *not* equal to the change in institutional ownership, which can be zero as long as institutional investors own the same proportion of shares at the beginning and at the end of the quarter. For stocks the distinction between the two is unimportant, but for ETFs it is critical because the number of shares outstanding changes frequently, and specifically from quarter to quarter.

significant. If liquidity continues to predict *IB* and *IS* over a quarterly horizon, it must be for reasons other than arbitrage activity.

A potential concern is that while institutional buying is unrestricted in the sense that any ETF can be bought, institutional selling is restricted to shares that are currently owned (as 13F only includes long positions). If liquidity determines institutional buying, then I may find a mechanical link between institutional selling (*IS*) and liquidity, because the only ETFs in the portfolio to sell are the liquid ones. However, I do not believe this to be an issue because the average (median) institutional ETF investors owns 9 (4) different ETFs at any given point in time during the 2006-2012 sample period.

### **2.3.7 Results: Institutional demand and liquidity**

The net demand by institutional investors is on average 1.99 percent of shares outstanding per quarter. Previously, I showed that liquidity is not an important determinant for quarterly net flows both in univariate tests (Table 2-1) and in regression analysis (Table 2-5). Similarly, for net institutional demand, there is no significant difference between funds with above median relative liquidity (2.09 percent of shares outstanding per quarter) vs. below (1.91 percent). Although it is not surprising to find that liquidity matters less over longer horizons, a calendar quarter is still a relatively short holding period, and I should not expect the significance to disappear altogether. However, short-term trading strategies require liquidity both for buys and for sells. Consistent with this idea, I find a significant difference in quarterly institutional buys and sells for ETFs with high vs. low relative liquidity (12.38 vs. 7.73 percent for buys and 10.44 vs. 5.83 percent for sells). The fact that liquidity is so important for institutional selling may explain why net institutional demand are not significantly related to ETF liquidity.

In order to study further the relationship between institutional demand (net, buys or sells) and liquidity, I estimate regressions similar to the fund flow regressions in Eq. (7). The results in Table 2-8 confirm the univariate results that liquidity does not matter for net institutional trades, but if I analyze institutional buys and sells separately, I find a positive and highly significant relationship. To illustrate the economic magnitude, an increase in ETF liquidity from the 25<sup>th</sup> to the 75<sup>th</sup> percentile is associated with a 2.7 percent increase in institutional buying (mean *IB* = 10.02 percent) and a 4.47 percent increase in institutional selling (mean *IS* = 8.21 percent).

[Table 2-8]

## 2.4 LIQUIDITY CLIENTELES AND SHORT-TERM TRADING

In this section, I provide further evidence on the existence of liquidity clienteles by first examining the importance of liquidity for short-term vs. long-term institutional ownership and second, by making a direct attempt to link liquidity to short-term institutional trading in ETFs using a measure that is unrelated to arbitrage activity.

### 2.4.1 Institutional ownership – baseline model

To arrive at a measure of total institutional ownership, I sum up the shares held by all 13-F investors for an ETF, and then scale this number by the median number of shares outstanding in the two weeks prior to the quarter-end. The median shares outstanding is used to mitigate the impact of outliers and timing errors in the reporting of shares held. This is a potential concern particularly because shares are on average created once every three days (see Table 2-3).

Aggarwal and Schofield (2012) suggest that institutions account for more than 80 percent of trading activity in ETFs. In contrast, our results show that the mean (median) institutional ownership is 41.93 (37.28) percent of shares outstanding (see Table 2-4, Panel B). Given that trading is costly, institutions should demand more liquidity than retail investors. The testable implication is that there is a positive relationship between institutional ownership and liquidity. As expected,  $IO$  is significantly higher for ETFs with high relative liquidity as opposed to low relative liquidity (50.81 vs. 33.22 percent at the mean, and the difference significant at 1 percent level). I also confirm this conjecture in the context of the following regression:

$$IO_{i,q} = a_0 + b_1 LIQ_{i,q-1} + b_2 Return_{i,q-1} + b_3 Fund_{i,q-1} + FE + \varepsilon_{i,t} \quad (12)$$

where  $LIQ_{i,q-1}$  = ETF & underlying portfolio or relative liquidity during quarter  $q-1$   
 $Fund_{i,q-1}$  = fund characteristics: expense ratios, age, AUM, tracking error and returns  
 $FE$  = style, sector and calendar time

Following Ferreira and Matos (2008), I include calendar time dummies to account for any residual correlation across ETFs in a given quarter (cross-sectional dependence) and cluster standard errors at the fund-level to account for residual correlations across quarters for a given fund (time-series dependence). An alternative solution is to use the Fama-MacBeth procedure to control for cross-sectional dependence in the residuals. All of our subsequent results remain robust to using this alternative methodology (unreported for conciseness). I also include style and

sector fixed to account for any time-invariant preferences that institutions may have in the aggregate (see e.g. Gompers and Metrick, 2001; Bennett, Sias and Stark, 2003).

[Table 2-9]

The results in Table 2-9 confirm that institutional ownership is higher among ETFs with higher liquidity, and ETFs with higher relative liquidity (significant at 1 percent for all measures), although underlying portfolio liquidity does not have a consistent sign (positive and significant for price impact; negative and significant for turnover). The ambiguity in the relationship between underlying portfolio liquidity and institutional ownership may, at least partly, reflect a lack of identification as I have already controlled for style, sector and calendar time fixed-effects. The only other consistent result is that *IO* is significantly higher for sector funds (by 8 to 18 percent) even after controlling for liquidity characteristics, possibly due to the popularity of sector rotation strategies among intuitional investors.

#### **2.4.2 Determinants of short- vs. long-term institutional ownership**

Although liquidity is beneficial to all investors, short-term investors should value liquidity the most (Amihud and Mendelson, 1986). In this section I therefore investigate the importance of liquidity for institutional ownership of ETFs by different types of institutions that are likely to differ in their ex-ante needs for liquidity. Our primary focus is on Bushee (1998)'s classification of institutional investors, which is based on factor and clustering analysis to classify institutional investors into three groups: transient (high portfolio turnover, diversified portfolios), quasi-indexer (low portfolio turnover, diversified portfolios) and dedicated (low portfolio turnover, concentrated investments). I view transient investors as short-term, and quasi-indexers as long-term. I drop the dedicated category from the analysis because their ownership of ETFs is on average less than one percent during our sample period.

A potential drawback with Bushee's classification is that it is applied at the portfolio level, which may falsely identify some institutions as long-term, when for our purposes they are short-term investors in ETFs. For instance, an institutional investor could have a significant passive allocation in core strategies (e.g. 75 percent buy-and-hold in a diversified portfolio of stocks) and a smaller allocation in securities used for short-term trading strategies (e.g. 25 percent of assets invested in highly liquid ETFs with a style or sector focus). For this reason, I also consider an

alternative classification based on three types of 13F institutions: i) commercial banks (BNK), ii) investment companies (INV) and iii) independent investment advisors (IIA). A commercial bank's trust department manages assets on behalf wealthy individual clients (trust accounts) and corporate pension plan clients. Investment companies are mainly mutual fund management companies; while investment advisors generally refer to investment banks, asset management companies, brokers and private wealth management companies. Section 13-F also contains several other types: insurance companies, public or private pension funds and endowments, and miscellaneous. I drop these additional types because they collectively own less than 6.5 percent of ETF assets compared with roughly 40 percent for BNK, INV and IIA combined. As discussed in the data section, I am unable to differentiate between INV and IIA, which is why I combine them together (INVA).

I conjecture that investment companies and advisors (INVA) are more likely to have shorter investment horizons relative to commercial banks (BNK) for three reasons. First, when I double sort on 13-F type (INVA and BNK) and investment horizon, I find that during our sample period, roughly 34.4 percent of INVA's can be classified as transient (based on of the fraction of 13-F securities held), whereas only 4.3 percent of BNK's can be classified as transient. Second, because INVA's include active investors such as mutual funds and hedge funds, active short-term trading strategies are more likely to be used by this group of investors as opposed to bank-trusts, which are more likely to be passive. Third, investment companies and advisors are also likely to require some liquidity within their holdings due to the open-end structure and performance sensitivity of many of their products. I therefore classify investment companies and independent investment managers as shorter-term and more active in their investment strategies, while commercial banks are classified as longer-term and more passive.

For our sample of ETFs, short-term (transient investors) own 16.56 percent of ETF shares on average. Is 16.56 percent high or low? To put things into perspective, I make a comparison between the proportion of transient investors in ETFs and the proportion of transient investors in the ETF's underlying basket of securities. In calculating the latter, I first compute the ownership by transient investors in each underlying security  $k$  held by ETF  $i$  in quarter  $q$ . I then calculate a dollar-weighted average for ETF  $i$ 's entire portfolio ( $IO_{TRA,i,q}^{UND} = \sum_{k=1}^N w_{i,k,q} IO_{TRA,q}^k$ ), scale this by the total institutional ownership ( $IO_{i,q}^{UND} = \sum_{k=1}^N w_{i,k,q} IO_q^k$ ) and compare the ratios:  $REL(IO_{TR}) = IO_{TRA,i,q}^{ETF} / IO_{i,q}^{ETF} - IO_{TRA,i,q}^{UND} / IO_{i,q}^{UND}$ . The full-sample mean proportion of transient investors is

considerable higher in ETFs relative to their underlying portfolios (38.0 percent versus 24.84 percent). Figure 2-2 depicts the evolution of the 25<sup>th</sup>, median and 75<sup>th</sup> percentile of the cross-sectional distribution of  $REL(IO_{TR})$ . The figure shows an upward trend with the median increasing from zero to above 10 percent.

[Figure 2-2]

Next, I analyze the importance of liquidity (and other fund characteristics) on institutional ownership by short-term (transient or INVA) vs. long-term (quasi-indexer or BNK) investors. I estimate regression (12) separately for each investor type, for the difference between short- and long-term institutional ownership and also for sub-samples of ETFs with a core and sector style classification. To conserve space, I combine the various liquidity measures together using the first principal component.

[Table 2-10]

The results in Table 2-10 reveal that while both short- and long-term investors are attracted to highly liquid ETFs, short-term investors care significantly more about liquidity. To illustrate the economic magnitudes, I compare difference in  $IO$  between funds at the 25<sup>th</sup> and 75<sup>th</sup> percentile of relative liquidity. The interquantile increase is 10.67 percent for short-term investors and 6.05 percent for long-term investors when I use the transient and quasi-indexer classifications; and by 10.13 and 3.03 percent when short- and long-term are based on INVA and BNK's. These differences are further amplified by a factor of two if I also take into consideration the fact that the average ownership by short-term investors is less than half of that for long-term investors. The differences between the impacts for short- vs. long-term investors are also statistically significant at the 1 percent level, and the differences are stronger for ETFs with a core style as opposed to sector ETFs. Similar effects are observed for ETF liquidity, while underlying portfolio liquidity is insignificant. The latter result is not surprising given that the signs on the individual measures of underlying portfolio liquidity are ambiguous (similar to Table 2-9); combining the various measures together using principal component analysis is therefore not expected to yield significant results.

The only other fund characteristic that affects short- vs. long-term investors differently is size (AUM), which enters with a negative and significant coefficient for short-term investors.

This may also reflect liquidity concerns. That is, short-term institutional ownership is larger in small funds because retail investors prefer bigger funds (unsophisticated investors may use size as a simple proxy for liquidity), while long-term investors are invariant.

### 2.4.3 Short-term trading and liquidity

Finally, I make a direct attempt to measure the importance of liquidity for institutional short-term trading in ETFs. Cremers and Pareek (2014) propose a measure of short-term trading that is based on institutional investors' average holding duration calculated from quarterly 13-F holdings as follows. For each ETF in a given institutional investor's portfolio, the duration measure is calculated by looking back in time to determine how long that particular ETF has been held continuously in that investor's portfolio, adjusted for intermediate buys and sells. Specifically:

$$Duration_{i,j,T-1} = \frac{(W-1)H_{i,j}}{H_{i,j} + B_{i,j}} + \sum_{t=T-W+1}^{T-1} \left( \frac{(T-t-1)\alpha_{i,j,t}}{H_{i,j} + B_{i,j}} \right) \quad (13)$$

- where  $B_{i,j}$  = total percentage of shares of ETF  $i$  bought by institution  $j$  between  $t=T-W$  and  $t=T-1$ ;  $t, T$  are in quarters  
 $H_{i,j}$  = percentage of total shares outstanding of ETF  $i$  held by institution  $j$  at time  $t=T-W$ .  
 $\alpha_{i,j,t}$  = percentage of total shares outstanding of ETF  $i$  bought or sold by institution  $j$  between time  $t-1$  and  $t$ , where  $\alpha_{i,j,t} > 0$  for buys and  $\alpha_{i,j,t} < 0$  for sells.  
 $W$  = maximum holding duration

The first term in (13) measures the duration of current holdings, while the second term in (13) is an adjustment term for any intermediate buying and selling due to, for instance, tax selling or window dressing. I choose a maximum holding duration ( $W$ ) of 12 quarters because any behavioural or informational effects beyond that horizon should be marginal. If ETF  $i$  is not held by institution  $j$  at time  $T-1$ , I set the duration measure to zero. To arrive at a measure of duration for ETF  $i$ , I calculate a weighted average of (13) across all institutions currently holding ETF  $i$  with the weights given by the ETF ownership of each institution. Before including institution  $j$  in the calculation, I require that  $j$  has existed for at least two years.

Descriptive statistics for ETF duration is provided in Table 2-1. The mean and median ETF duration are very close to each other in the overall sample (2.84 and 2.77 quarters). ETF duration

is slightly lower for value and growth funds relative to blend funds, and lower for sector funds relative to core funds. Do these numbers imply that ETFs are used for short-term trading? To answer this question, I make a comparison between ETF and underlying portfolio duration. For each ETF  $i$ , the underlying portfolio duration is calculated as the dollar-weighted average duration of each stock in the portfolio. Table 2-1 shows that ETF duration is on average 2.01 quarters *less* than the corresponding underlying portfolio duration, which confirms that there is more short-term trading in ETFs.

A unique feature of ETF duration is that it ignores all intra-quarterly trading. As a result, ETF duration measures short-term trading unrelated to arbitrage trades. The reason is that mispricing is very short-lived: a regression of the daily ETF premium (8) on its lags shows that only the first few daily lags are statistically significant. If liquidity continues to be an important determinant of short-term trading it must be because liquidity is attractive for short-term trading in response to non-fundamental demand.

One potential drawback with ETF duration is that it is mechanically related to the average age of the fund's investor base. Specifically, the institution specific holding duration (13) starts at zero for an investor's first investment in the ETF, whether by existing or new ETF investors, and can only reach the maximum holding duration 12 quarters later. This is a particular concern for us because ETFs are a relatively recent investment vehicle for many institutional investors, and because many new ETFs enter throughout the sample period. To address this concern, I control for the time (in quarters) since ETF  $i$  was first purchased by investor  $j$ , averaged across all investors with the weights given by the ownership of ETF  $i$  by investor  $j$ . I also control for fund age in all regressions where ETF duration is the dependent variable. In addition, I require that ETF  $i$  has existed for at least two years before including it in the sample.

Next, I investigate the relationship between ETF duration and liquidity, controlling for fund characteristics, returns and fixed effects. The explanatory variables are the same as those in the institutional ownership regression (12), except that I now also add fund age and the age of its investor base.

[Table 2-11]

Table 2-11 presents the results. Specification (1) – with only style fixed effects – shows that ETF liquidity is negatively related to ETF duration ( $t = 9.77$ ), while underlying portfolio liquidity is

positively related ( $t = 2.62$ ). This implies that highly liquid ETFs that invest in securities with lower liquidity have more short-term trading, which I confirm in specification (2) where I only include relative liquidity. The economic impact of an increase in relative liquidity from the 25<sup>th</sup> percentile to the 75<sup>th</sup> percentile is associated with a decrease in ETF duration by 0.64 quarters, which is meaningful when compared to the average ETF duration of 2.84 quarters.

Controlling for time fixed effects (specification (3) and (4)) does not alter the significance of ETF liquidity (3) or relative liquidity (4) from before. However, underlying portfolio liquidity is now insignificant, which is not surprising because time fixed effects soak up much of the variation in underlying portfolio liquidity. Even after I control for time and ETF fixed effects (specifications (5) and (6)), ETF liquidity and relative ETF liquidity remain significant ( $t = 3.47$  and 4.65 respectively). The only other fund characteristics that remain significant across all specifications are size and the age of the funds investor base: bigger funds and fund with an old investor base are on average held for longer durations.

Overall, the results in this section show that funds with higher liquidity have more short-term trading (i.e. lower average institutional holding periods) for reasons unrelated to arbitrage trades.

## **2.5 SUMMARY AND CONCLUSION**

Exchange-Traded Funds provide investors with easy access to popular investment styles (e.g. large, small, value, growth and sector) at a cost that is on average lower relative to their underlying basket of securities. Apart from being exchange-traded with real-time prices, a unique feature of ETFs is their open-end structure via the share creation/redemption process, which adds a second layer of liquidity on top of secondary market trading volume.

I started with a quote from Vanguard's founder, John Bogle, that because ETFs are listed on an exchange, they lure investors to trade them too often. In this paper, I argue that ETFs that are relatively liquid (in four different dimensions) indeed attract a clientele of short-term investors, and more explicitly, short-term trading in ETFs. First, I document that liquidity is an important determinant of weekly or monthly net fund flows after controlling for a host of fund and macroeconomics characteristics, while quarterly flows are much more responsive to trading costs over longer horizons (namely expense ratios). Second, I show that measures of institutional buying and selling are both significantly higher in more liquid ETFs consistent with the idea that

short-term trading requires liquidity both for buys and sells. Third, I document that the proportion of short-term investors in ETFs is more than 10 percent higher relative to their underlying baskets and that liquidity is significantly more important for short-term institutional investors relative to long-term (by two to six times). Third, by using a direct measure of short-term trading based on average quarterly institutional holding periods, I show that there is more short-term trading in ETFs relative to their underlying portfolios. Moreover, ETF holding periods are significantly lower for more liquid ETFs for reasons unrelated to arbitrage activity.

One important implication of our findings is that as the ETF market continues to grow, it will likely facilitate a higher volume of short-term trading, thereby having an impact on market efficiency and volatility, as well as causing higher correlations in the underlying basket of securities.<sup>19</sup>

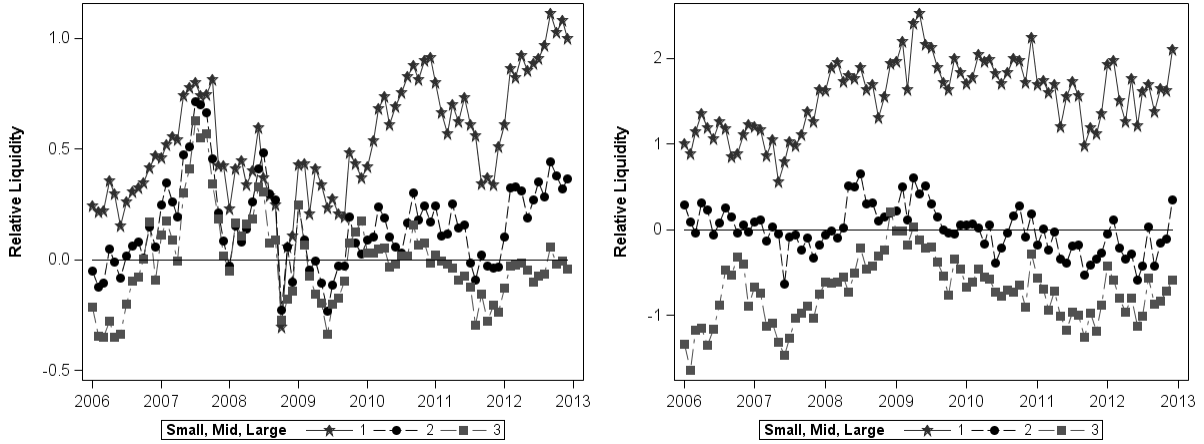
---

<sup>19</sup> For recent discussion and evidence of ETFs causing higher correlations in the underlying basket, see Da and Shive (2013), and Katherine Fogertey and Robert D. Boroujerdi (2015)'s Goldman Sachs study, "ETFs: the Rise of the Machines".

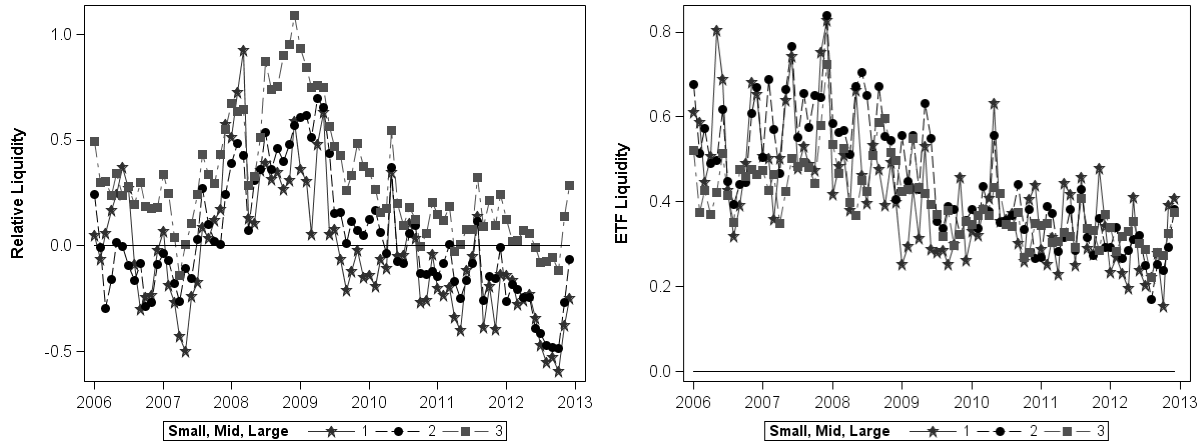
**Figure 2- 1: Time-series variability of relative liquidity (Balanced Panel)**

The figures below illustrate the cross-sectional mean relative liquidity for ETFs with the Morningstar style classification small-, mid- and large-cap.

**Panel A: Relative liquidity via spreads (LEFT) and price impact (RIGHT)**

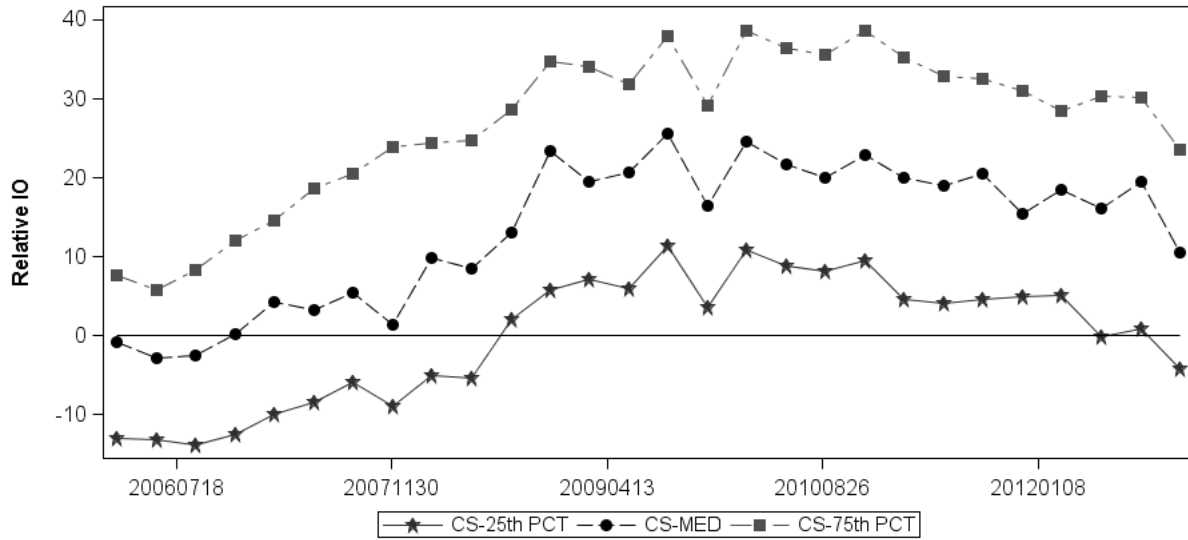


**Panel B: Relative liquidity via turnover (LEFT) and share creation activity (RIGHT)**



**Figure 2- 2: Time-series Variability of Relative Institutional Ownership**

The figures below illustrate the cross-sectional median, 25<sup>th</sup> and 75<sup>th</sup> percentile of the relative institutional ownership of transient (short-term) investors, calculated as:  $REL(IO)_{TRA,q} = (IO_{TRA,q}^{ETF}/IO_q^{ETF}) - (IO_{TRA,q}^{UND}/IO_q^{UND})$ .



**Table 2- 1: Descriptive statistics**

Means and medians are reported for the following variables: Assets Under Management (AUM), net expense ratios, ETF duration and relative ETF duration. All variables are measures over the entire sample period, from January 2006 to December 2012. ETF duration refers to the average quarterly institutional holding duration for the ETF, while REL(duration) refers to the difference between ETF duration and the dollar-weighted duration of an ETF's underlying portfolio of stocks. Descriptive statistics are reported by Morningstar's 3-by-3 style classification for funds with a core style, and by Morningstar sector for sector funds.

Style	# (2012)	AUM (\$billion)		Exp. Ratio (%)		ETF Duration		REL (Duration)	
		Mean	MED	Mean	MED	Mean	MED	Mean	MED
Large – Value	12	2.21	0.37	0.26	0.20	3.07	3.00	-2.42	-2.62
Large – Blend	21	7.54	0.81	0.20	0.20	3.27	3.14	-2.02	-2.14
Large – Growth	12	3.31	0.34	0.25	0.20	3.01	2.92	-1.79	-1.81
Mid – Value	5	1.11	0.63	0.25	0.25	3.58	3.60	-1.02	-0.85
Mid – Blend	6	2.60	1.16	0.22	0.20	3.67	3.70	-0.54	-0.59
Mid – Growth	5	0.97	0.46	0.26	0.25	3.24	3.23	-0.72	-0.59
Small – Value	3	1.19	0.62	0.25	0.25	3.18	3.36	-1.20	-1.19
Small – Blend	10	2.54	0.70	0.26	0.20	3.01	3.07	-1.17	-1.11
Small – Growth	6	1.41	0.73	0.25	0.25	2.99	3.04	-0.86	-0.82
Communication	2	0.39	0.38	0.41	0.48	3.39	3.32	-1.37	-1.59
Cons. Cyclical	7	0.48	0.27	0.37	0.35	2.07	1.96	-2.46	-2.74
Cons. Defens.	4	0.84	0.36	0.37	0.28	2.63	2.86	-2.77	-2.73
Energy	9	1.23	0.55	0.40	0.47	2.69	2.63	-1.81	-1.86
Financial	13	0.71	0.19	0.42	0.47	2.05	2.01	-3.00	-3.01
Real Estate	10	1.22	0.34	0.40	0.48	2.55	2.37	-2.72	-2.69
Health Care	11	0.70	0.40	0.42	0.47	2.54	2.48	-2.22	-2.15
Industrials	8	0.46	0.24	0.45	0.48	2.39	2.34	-2.67	-2.67
Materials	5	0.62	0.49	0.37	0.35	2.12	2.10	-2.56	-2.48
Technology	12	0.62	0.21	0.45	0.50	2.68	2.56	-2.04	-2.11
Utilities	4	1.19	0.64	0.36	0.27	2.99	2.89	-2.38	-2.51
All	165	2.24	0.40	0.33	0.25	2.84	2.77	-2.01	-2.10

**Table 2- 2: Correlations among liquidity variables**

This table reports correlations among our liquidity measures (ETF liquidity versus relative liquidity) for the full sample of ETFs between January 2006 and December 2012: share creation activity (CREATION), quoted spreads (QSPR), Amihud's price impact (AMIHU), and turnover (TO).

<b>Variables</b>	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>	<b>(7)</b>
<b>Panel A: Liquidity measures</b>							
(1) CREATION	1.00						
(2) ETF QSPR	0.04	1.00					
(3) ETF AMIHU	0.36	0.68	1.00				
(4) ETF TO	0.68	0.17	0.65	1.00			
(5) REL (QSPR)	0.19	0.65	0.67	0.39	1.00		
(6) REL (AMIHU)	0.34	0.47	0.83	0.59	0.77	1.00	
(7) REL (TO)	0.66	0.27	0.70	0.95	0.41	0.62	1.00
<b>Panel B: Liquidity shocks (based on an AR(1) model)</b>							
(1) CREATION	1.00						
(2) ETF QSPR	-0.08	1.00					
(3) ETF AMIHU	0.17	0.23	1.00				
(4) ETF TO	0.30	-0.12	0.45	1.00			
(5) REL (QSPR)	-0.02	0.76	0.15	0.09	1.00		
(6) REL (AMIHU)	0.21	0.08	0.87	0.57	0.13	1.00	
(7) REL (TO)	0.30	0.01	0.57	0.91	0.11	0.65	1.00

**Table 2- 3: ETF share creation activity**

This table reports summary statistics for the percentage of days with share creation/redemption activity and the size (absolute value) of share creations/redemptions conditional on activity days. The high and low categories of REL(LIQ) refer to above or below median relative liquidity, which is calculated the average rank across the all four relative liquidity measures (spreads, price impact, turnover and share creation activity). REL(LIQ) is measured in the month prior to the share creation activity. There are also 3 size (AUM) categories, and 2 liquidity categories (High and Low). In calculating the % of days with share creation activity, I first calculate the proportion for each ETF, and then average across all ETFs in the cross-section. Shaded numbers (in gray) denote statistical significance at the 5 % level. The Wilcoxon Signed Ranks test is used under the null of equality in medians between ETFs with high vs. low relative liquidity.

Category	Conditional on activity days							
	% of days		Value (in \$million)			% of AUM (= % of SHARES)		
	AVG	MED	AVG	MED	STD	AVG	MED	STD
All	30.9	22.7	69.6	12.4	292.5	2.8	1.0	9.9
<b>AUM –Top tercile</b>								
Top	57.9	55.8	99.5	20.5	355.1	1.4	0.6	3.5
REL LIQ High	59.3	58.0	104.0	22.3	361.3	1.5	0.6	3.1
REL LIQ Low	34.6	29.4	19.0	6.6	74.5	1.0	0.4	6.6
<b>AUM – Middle tercile</b>								
Middle	29.4	23.8	16.7	7.4	30.0	3.5	1.7	10.8
REL LIQ High	34.6	31.3	20.0	9.5	33.2	4.1	2.0	12.5
REL LIQ Low	20.2	18.2	9.3	5.6	19.9	2.4	1.3	5.2
<b>AUM –Bottom tercile</b>								
Bottom	12.9	9.3	7.0	3.8	11.4	8.8	4.3	18.5
REL LIQ High	17.7	14.3	10.4	5.9	14.8	9.5	4.8	20.2
REL LIQ Low	10.4	7.5	5.6	3.2	9.1	8.5	4.2	17.7

**Table 2- 4: Measures of ETF demand and liquidity**

This table reports summary statistics for different measures of ETF demand: net fund flows in Panel A, and measures of institutional demand and ownership in Panel B. The high and low categories of REL(LIQ) refer to above or below median relative liquidity, which is calculated the average rank across the all four relative liquidity measures (spreads, price impact, turnover and share creation activity). REL(LIQ) is measured in the period prior to measuring ETF demand.

<b>PANEL A: Net fund flow (in percent) by horizon (t)</b>									
<b>Category</b>	<b>t = Week</b>			<b>t = Month</b>			<b>t = Quarter</b>		
	<b>Mean</b>	<b>Med</b>	<b>STD</b>	<b>Mean</b>	<b>Med</b>	<b>STD</b>	<b>Mean</b>	<b>Med</b>	<b>STD</b>
All funds	0.48	0.00	5.23	2.15	0.00	11.54	2.70	2.04	19.58
High REL(LIQ), t-1	0.55	0.00	6.23	2.13	0.67	12.44	2.14	2.37	20.17
Low REL(LIQ), t-1	0.40	0.00	3.98	2.18	0.00	10.62	3.25	1.46	19.02
<b>AUM – Top tercile</b>									
Top	0.24	0.05	4.45	1.03	0.68	7.62	1.43	1.98	13.40
High REL(LIQ), t-1	0.22	0.06	4.76	0.99	0.63	7.96	1.08	1.90	13.70
Low REL(LIQ), t-1	0.34	0.02	2.05	1.25	0.92	5.20	3.30	2.60	11.53
<b>AUM – Middle tercile</b>									
Top	0.46	0.00	5.36	1.90	0.49	11.29	2.20	2.45	18.74
High REL(LIQ), t-1	0.66	0.00	7.10	2.42	0.92	14.18	1.77	3.01	22.44
Low REL(LIQ), t-1	0.29	0.00	2.96	1.44	0.03	7.76	2.60	2.16	14.60
<b>AUM – Bottom tercile</b>									
Top	0.72	0.00	5.77	3.51	0.00	14.53	4.39	1.22	24.66
High REL(LIQ), t-1	1.77	0.00	8.87	6.58	0.07	20.52	7.41	9.26	32.34
Low REL(LIQ), t-1	0.48	0.00	4.76	2.83	0.00	12.74	3.66	0.07	22.39

<b>PANEL B: Institutional demand and ownership</b>								
<b>Category</b>	<b>Net Institutional trades</b>		<b>Institutional buys</b>		<b>Institutional sells</b>		<b>Institutional ownership</b>	
	<b>Mean</b>	<b>Med</b>	<b>Mean</b>	<b>Med</b>	<b>Mean</b>	<b>Med</b>	<b>Mean</b>	<b>Med</b>
All funds	1.99	0.74	10.02	5.61	8.21	4.92	41.93	37.28
High REL(LIQ), t-1	2.09	0.92	12.38	7.95	10.44	7.05	50.81	46.15
Low REL(LIQ), t-1	1.91	0.69	7.73	4.14	5.83	3.43	33.22	31.13

**Table 2- 5: Explaining ETF inflows – base model**

This table reports a regression of monthly net fund flows on various fund characteristics and macro variables (all lagged), including inflows, returns, monthly average premium, AUM, age, tracking error, Pastor and Stambaugh (2003) aggregate level of liquidity, VIX, Term-Spread and style returns (own and distant). All specifications include time (month, year), sector and style (based on Morningstar's 3x3 size-valuation matrix) fixed effects. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 percent level. Standard errors are clustered by time (t-statistics in brackets).

Variables	All funds		Core styles	Sector styles
	(1)	(2)	(3)	(4)
<i>Fund characteristics</i>				
Inflows, t-1	-0.022 (0.91)	-0.066*** (2.81)	-0.061 (1.20)	-0.080*** (3.54)
Inflows, t-2	0.038** (2.55)	0.010 (0.73)	0.059** (2.61)	-0.017 (0.95)
Inflows, t-3	0.040** (2.64)	0.017 (1.09)	0.082*** (3.32)	-0.009 (0.46)
ETF Return, t-1	0.196*** (4.45)	0.296*** (6.58)	0.184*** (4.75)	0.359*** (6.78)
ETF Return, t-2	0.016 (0.54)	0.085** (2.46)	0.052 (1.37)	0.113*** (2.87)
ETF Return, t-3	0.047 (1.45)	0.084** (2.41)	0.036 (1.00)	0.123*** (3.10)
Cumulative inflows, t-4		0.058 (0.36)	-0.219 (1.35)	0.236 (0.79)
Expense Ratio, t-1		-3.566*** (3.32)	-3.862** (2.52)	-9.351*** (5.98)
Age, t		-1.306*** (8.00)	-0.960*** (5.47)	-1.344*** (3.97)
Age^2, t		0.072*** (8.06)	0.042*** (5.10)	0.100*** (4.88)
Average premium ( <i>PREM</i> ), t-1		24.682*** (5.24)	25.703*** (4.10)	22.764*** (3.23)
Assets ( <i>AUM</i> ), t-3		-0.458*** (4.44)	-0.080 (0.69)	-1.232*** (5.27)
Tracking error, t-1		3.225 (1.11)	2.634 (0.76)	1.855 (0.64)
<i>Controls – Macro</i>				
Own-Distant style return, t-1		0.160*** (4.08)	0.225*** (6.26)	0.120* (1.87)
Own-Distant style return, t-2		0.025 (0.61)	0.066* (1.93)	-0.021 (0.29)
Own-Distant style return, t-3		-0.022 (0.72)	0.020 (0.66)	-0.074 (1.43)
VIX index, t-1		0.145*** (2.69)	0.146*** (3.28)	0.138* (1.89)
Term Spread, t-1		1.261** (2.45)	2.011*** (3.46)	0.725 (1.23)
P&S (2003) liquidity, t-1		-2.162 (0.65)	1.064 (0.34)	-5.124 (1.22)
$R^2$	0.030	0.071	0.127	0.078
<i>N</i>	12,194	11,930	5,785	6,145

**Table 2- 6: Inflows and liquidity**

This table reports the results from regressions of monthly ETF inflows on liquidity characteristics; fund characteristics and control variables. I use share creation/redemption activity (Create), quoted spreads, price impact and turnover to measure liquidity (ETF, underlying and relative). All liquidity measures are signed such that positive numbers indicate liquidity. Panel A includes liquidity levels; Panel B uses shocks to liquidity. Control variables includes three lags of monthly flows and returns, cumulative inflows from 2006, Pastor and Stambaugh (2003) aggregate level of liquidity, VIX index, term-spread and four lags of Morningstar 3x3 style returns (own and distant). All specifications include time (year, month), sector and style fixed effects. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 percent level. Standard errors are clustered by calendar time (t-statistics in brackets).

Variables	Create	Quoted Spread	Price Impact	Turnover			
<b>Panel A: Liquidity levels</b>							
Exp. Ratio, t-1	-6.691*** (5.45)	-6.567*** (5.05)	-6.596*** (5.31)	-7.032*** (5.70)	-7.289*** (5.95)	-6.999*** (5.72)	-6.966*** (5.70)
PREM, t-1	27.487*** (5.87)	26.411*** (5.63)	26.439*** (5.63)	26.717*** (5.67)	27.116*** (5.78)	27.127*** (5.78)	27.034*** (5.75)
AUM, t-1	-0.524*** (4.97)	-0.547*** (3.43)	-0.539*** (3.95)	-1.733*** (5.73)	-1.118*** (6.74)	-0.739*** (5.79)	-0.732*** (5.78)
Track. Err., t-1	1.662 (0.58)	2.588 (0.82)	2.481 (0.87)	1.145 (0.42)	1.027 (0.37)	1.267 (0.47)	1.654 (0.60)
ETF LIQ, t-1		0.678* (1.72)		1.058*** (4.82)		1.086*** (5.05)	
UND LIQ, t-1		-0.586 (1.20)		-0.246** (2.20)		-0.448 (1.01)	
REL LIQ, t-1	1.778*** (4.48)		0.644** (2.33)		0.587*** (5.50)		1.069*** (5.03)
Controls: Dummies	YES	YES	YES	YES	YES	YES	YES
Style, sector	YES	YES	YES	YES	YES	YES	YES
Year, month	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.080	0.075	0.075	0.082	0.079	0.082	0.082
N	11,930	11,928	11,928	11,929	11,929	11,930	11,930
<b>Panel B: Liquidity shocks</b>							
Exp. Ratio, t-1	-4.159 (0.83)	-4.528 (0.91)	-4.413 (0.89)	-4.239 (0.86)	-4.565 (0.93)	-3.665 (0.73)	-3.385 (0.67)
PREM, t-1	25.468*** (5.33)	24.746*** (5.52)	25.562*** (5.39)	24.587*** (5.22)	24.531*** (5.21)	24.856*** (5.60)	24.324*** (5.24)
AUM, t-1	-2.469*** (2.78)	-2.562*** (2.86)	-2.545*** (2.83)	-2.688*** (3.00)	-2.663*** (2.98)	-2.470*** (2.84)	-2.485*** (2.81)
Track. Err., t-1	-2.963 (1.14)	-3.436 (1.27)	-2.671 (1.05)	-2.513 (0.97)	-2.744 (1.07)	-3.193 (1.32)	-2.319 (0.92)
ΔETF LIQ, t-1		-0.195 (0.33)		1.100*** (4.33)		2.731*** (6.62)	
ΔUND LIQ, t-1		-1.315 (1.30)		-0.254 (0.31)		0.405 (0.24)	
ΔREL LIQ, t-1	1.796** (2.56)		0.090 (0.15)		1.106*** (4.38)		2.860*** (6.89)
Controls: Dummies	YES	YES	YES	YES	YES	YES	YES
ETF	YES	YES	YES	YES	YES	YES	YES
Year, month	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.118	0.117	0.116	0.118	0.118	0.124	0.124
N	11,920	11,916	11,916	11,918	11,918	11,920	11,920

**Table 2- 7: Horizon of net flows and liquidity**

This table reports regressions of ETF net flows (measured over weekly, monthly or quarterly horizons) on liquidity characteristics; fund characteristics and control variables (see Table 5 for more details). I use share creation/redemption activity (Create), quoted spreads, price impact and turnover to measure liquidity (ETF, underlying and relative). All liquidity measures are signed such that positive numbers indicate liquidity. Panel A includes liquidity levels; Panel B uses shocks to liquidity. All specifications include time (year, month), sector and style fixed effects. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 percent level. Standard errors are clustered by calendar time (t-statistics in brackets).

Variables	Weekly			Monthly			Quarterly		
	All	Sector	Core	All	Sector	Core	All	Sector	Core
<b>PANEL A: Liquidity levels</b>									
EXP, t-1	0.729 (0.72)	2.346 (1.47)	-0.281 (0.19)	-0.438 (0.08)	8.179 (1.00)	-4.724 (0.72)	-36.51*** (3.19)	-15.92 (0.94)	-63.36** (2.42)
Age, t	1.252 (0.85)	-3.108 (1.44)	5.597*** (3.95)	136.196 (1.56)	154.686 (1.40)	109.028 (1.43)	-45.296 (0.17)	-99.387 (0.21)	-38.427 (0.14)
Age^2, t	-0.001 (0.36)	0.001 (0.26)	0.001 (0.28)	0.011 (0.78)	0.036 (1.51)	-0.001 (0.07)	0.010 (0.23)	0.051 (0.72)	-0.020 (0.41)
PREM, t-1	6.802*** (7.84)	7.785*** (7.17)	4.819*** (4.63)	25.121*** (5.33)	24.233*** (3.56)	24.248*** (3.90)	74.496*** (3.91)	90.163*** (2.95)	53.047** (2.63)
AUM, t-1	-0.492** (2.18)	-0.771*** (3.05)	-0.213 (0.61)	-2.965*** (3.43)	-3.579*** (3.65)	-2.757* (1.69)	-7.613*** (4.69)	-6.181*** (2.84)	-9.779*** (4.23)
Track. Err., t-1	0.291 (0.70)	0.320 (0.66)	0.315 (0.62)	-2.191 (0.81)	-1.993 (0.74)	1.060 (0.27)	-11.377 (1.41)	-17.170* (1.91)	0.950 (0.07)
ETF LIQ, t-1	1.055*** (12.22)	1.551*** (12.88)	0.300*** (2.92)	3.047*** (6.67)	4.119*** (6.23)	1.636*** (3.06)	0.748 (0.54)	0.526 (0.26)	2.376 (1.36)
UND LIQ, t-1	0.053 (0.41)	0.080 (0.44)	0.051 (0.34)	-0.019 (0.03)	0.336 (0.39)	-0.536 (0.70)	-3.185* (1.67)	-4.006 (1.23)	-1.279 (0.65)
Controls:	YES	YES	YES	YES	YES	YES	YES	YES	YES
Dummies									
ETF	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year, month	YES	YES	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.054	0.059	0.063	0.125	0.125	0.171	0.194	0.180	0.293
N	51,699	26,458	25,241	11,927	6,144	5,783	3,832	1,958	1,874
<b>PANEL B: Liquidity shocks</b>									
EXP, t-1	-0.264 (0.26)	0.240 (0.15)	-0.216 (0.15)	-10.638** (2.01)	-6.010 (0.74)	-13.750** (2.03)	-25.552** (2.32)	-8.210 (0.42)	-35.409** (2.26)
Age, t	1.032 (0.71)	-3.250 (1.49)	5.429*** (3.92)	145.516* (1.76)	137.024 (1.33)	122.260 (1.59)	-214.249 (0.71)	-497.990 (0.95)	-42.856 (0.18)
Age^2, t	0.002 (0.82)	0.004 (0.86)	0.002 (0.61)	0.030** (2.22)	0.062*** (2.73)	0.005 (0.35)	0.026 (0.53)	0.097 (1.27)	-0.037 (0.70)
PREM, t-1	6.781*** (7.62)	7.739*** (7.21)	4.819*** (4.51)	26.194*** (5.93)	25.711*** (3.83)	24.931*** (4.06)	70.074*** (3.63)	79.464** (2.58)	54.082*** (2.68)
AUM, t-1	-0.307 (1.36)	-0.417* (1.67)	-0.221 (0.66)	-4.682*** (10.37)	-4.853*** (8.63)	-3.985*** (6.62)	-10.344*** (7.95)	-9.397*** (5.20)	-11.636*** (7.48)
Track. Err., t-1	0.041 (0.10)	0.047 (0.10)	0.224 (0.44)	-1.743 (0.67)	-1.557 (0.63)	1.010 (0.24)	-5.839 (0.73)	-9.282 (0.98)	-1.448 (0.12)
ΔETF LIQ, t-1	0.336*** (10.23)	0.482*** (10.19)	0.136*** (3.88)	0.898*** (5.84)	1.229*** (5.30)	0.511*** (3.16)	-0.134 (0.33)	0.183 (0.29)	-0.020 (0.04)
ΔUND LIQ, t-1	0.011 (0.24)	-0.007 (0.12)	0.039 (0.77)	-0.128 (0.63)	-0.106 (0.42)	-0.125 (0.57)	0.150 (0.22)	0.326 (0.35)	-0.305 (0.34)
Controls:	YES	YES	YES	YES	YES	YES	YES	YES	YES
Dummies									
ETF	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year, month	YES	YES	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.051	0.052	0.064	0.121	0.122	0.163	0.186	0.174	0.286
N	51,640	26,430	25,210	11,962	6,171	5,791	3,996	2,076	1,920

**Table 2- 8: Institutional trading and liquidity**

This table reports the results for regressions of quarterly net institutional trades, institutional buys or institutional sells on measures of liquidity (ETF and underlying portfolio) and fund characteristics. ETF (underlying portfolio) liquidity is calculated as the first principal component of ETF (underlying portfolio) across all liquidity measures. For a more detailed description of the fund specific variables, see Table 5. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 percent level. Standard errors are clustered by calendar time (*t*-statistics in brackets).

Variables	Net trades	Institutional buying			Institutional selling		
	Full	Full	Core	Sector	Full	Core	Sector
Y., q-1	-0.211*** (7.04)	-0.023 (0.62)	0.065 (1.16)	-0.049 (1.22)	-0.095*** (3.82)	-0.022 (0.46)	-0.134*** (4.64)
Y., q-2	-0.078** (2.66)	0.026 (0.69)	0.031 (0.53)	0.019 (0.51)	-0.052* (2.03)	0.017 (0.40)	-0.085** (2.51)
ETF Ret, q-1	0.228*** (6.68)	0.189*** (7.47)	0.195*** (7.27)	0.193*** (5.70)	-0.026 (1.10)	-0.035 (0.99)	-0.023 (0.84)
Exp. Ratio, q-1	-0.470 (0.05)	-0.829 (0.09)	7.328 (0.74)	0.995 (0.06)	-6.187 (0.85)	-7.157 (0.77)	-11.791 (1.69)
Age, q	417.151** (2.29)	-44.888 (0.18)	455.683** (2.66)	-529.470 (1.36)	-430.777 (1.51)	-124.801 (0.67)	-742.492* (1.96)
Age^2, q	-0.045 (1.18)	-0.062 (1.58)	-0.028 (0.79)	-0.084 (1.60)	-0.016 (0.59)	-0.005 (0.19)	-0.018 (0.49)
PREM, q-1	30.669** (2.24)	-7.004 (0.56)	-8.371 (0.61)	-10.021 (0.47)	-24.615*** (3.89)	-19.672* (1.77)	-30.286*** (3.74)
AUM, q-1	-4.110*** (4.51)	-4.280*** (4.31)	-1.662** (2.31)	-5.665*** (3.99)	0.103 (0.18)	0.071 (0.11)	-0.020 (0.03)
Track Err., q-1	-12.118** (2.60)	0.283 (0.06)	8.058 (1.01)	-2.664 (0.47)	8.066** (2.54)	9.674 (1.19)	5.940 (1.54)
ETF LIQ, q-1	0.344 (0.36)	2.592*** (3.19)	3.096*** (2.90)	2.824* (2.05)	4.297*** (7.24)	2.644*** (4.32)	6.174*** (7.31)
UND LIQ, q-1	0.950 (0.91)	0.469 (0.42)	1.289 (0.98)	0.322 (0.22)	0.328 (0.43)	0.685 (0.75)	0.176 (0.16)
Dummies:							
ETF	YES	YES	YES	YES	YES	YES	YES
Calendar	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.139	0.372	0.364	0.338	0.452	0.392	0.422
N	3,531	2,984	1,392	1,592	2,964	1,382	1,582

**Table 2- 9: Institutional ownership and ETF Liquidity**

This table reports the results for regressions of an ETF's institutional ownership on various dimensions of ETF Liquidity; share creation activity, quoted spreads, Amihud's price impact or turnover; and other ETF characteristics (see Table 5 for additional details on the variable abbreviations). All specifications include style, sector and year-month fixed effects. \*/\*\*/\*\*\*\* denotes statistical significance at the 10, 5 and 1 percent level. Standard errors are clustered by fund (*t*-statistics in brackets).

Variables	Creation	Quoted Spread	Price Impact	Turnover			
Exp. Ratio, q-1	4.842 (0.55)	7.840 (0.66)	5.120 (0.43)	3.485 (0.44)	0.304 (0.03)	1.223 (0.16)	1.382 (0.17)
Age, q	0.639 (0.57)	-0.923 (0.79)	-1.521 (1.33)	1.586 (1.44)	-1.463 (1.23)	1.562 (1.36)	1.621 (1.40)
Age <sup>2</sup> , q	0.015 (0.24)	0.065 (1.02)	0.118** (1.99)	-0.123* (1.82)	0.121** (1.99)	-0.099 (1.44)	-0.103 (1.48)
AUM, q-1	2.643*** (2.83)	2.325** (2.16)	3.226*** (2.99)	-9.822*** (6.25)	0.279 (0.16)	0.863 (0.94)	0.887 (0.97)
Sector dummy	9.989*** (4.05)	18.659*** (5.65)	18.425*** (5.44)	9.392*** (3.65)	15.161*** (4.41)	6.707** (2.42)	6.808** (2.44)
ETF Ret, q-1	0.026 (0.64)	0.047 (0.93)	0.035 (0.71)	0.180*** (3.99)	0.048 (1.00)	0.077* (1.87)	0.065 (1.60)
Track Err., q-1	-4.736 (0.55)	22.124** (1.97)	9.880 (0.93)	0.235 (0.03)	1.310 (0.13)	-4.334 (0.47)	-0.487 (0.05)
ETF LIQ, q-1		11.620*** (3.19)		11.338*** (9.37)		11.345*** (9.04)	
UND LIQ, q-1		0.585 (0.16)		1.993** (2.43)		-6.585** (2.23)	
REL LIQ, q-1	23.276*** (9.87)		7.231** (2.37)		3.452*** (3.02)		11.270*** (9.02)
Dummies							
Style	YES	YES	YES	YES	YES	YES	YES
Calendar	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.432	0.287	0.276	0.464	0.297	0.447	0.444
N	3,631	3,610	3,610	3,610	3,610	3,610	3,610

**Table 2- 10: Determinants of short- vs. long-term institutional ownership**

This table reports results for regressions of ETF institutional ownership (by short-, long- and short-minus-long-term investors) on liquidity (ETF, underlying portfolio and relative liquidity) as well as fund characteristics. ETF (underlying portfolio) liquidity is calculated as the first principal component of ETF (underlying portfolio) share creation activity, spreads, price impact and turnover. Panel A is estimated for institutions classified as transient (short-term) and quasi-indexers (long-term), while Panel B is estimated for institutions classified as investment companies and independent investment advisors (short-term and active) vs. bank trusts (long-term and passive). All specifications include calendar, sector and style fixed effects. \*/\*\*/\*\*\*/ denotes statistical significance at the 10, 5 and 1 percent level. Standard errors are clustered by fund (t-statistics in brackets).

Variables	All funds			Core styles			Sector styles		
	Short	Long	ΔS-L	Short	Long	ΔS-L	Short	Long	ΔS-L
<b>PANEL A: Transient (short-term) vs. Quasi-indexers (long-term)</b>									
Exp., q-1	10.156** (2.55)	-4.619 (0.74)	16.751*** (2.61)	13.588*** (2.83)	13.536 (1.61)	0.634 (0.08)	4.494 (0.78)	-18.550** (2.47)	26.767*** (3.20)
Age, q	1.576** (2.10)	0.393 (0.58)	1.411 (1.58)	1.018 (1.15)	-0.806 (0.70)	1.814* (1.82)	2.664*** (3.16)	1.203 (1.21)	1.665 (1.24)
Age^2, q	-0.075 (1.43)	-0.045 (1.22)	-0.048 (0.79)	-0.038 (0.60)	0.008 (0.12)	-0.049 (0.84)	-0.113** (2.11)	-0.056 (0.88)	-0.082 (0.98)
AUM, q-1	-3.512*** (7.25)	1.356* (1.76)	-4.589*** (5.72)	-2.112*** (3.81)	3.835*** (2.84)	-5.867*** (4.89)	-4.855*** (6.06)	-0.384 (0.44)	-3.890*** (3.57)
ETF Ret, q-1	0.069** (2.10)	0.047 (1.61)	0.014 (0.29)	0.147** (2.49)	-0.050 (0.79)	0.197*** (2.84)	0.093** (2.47)	0.078** (2.42)	-0.004 (0.07)
Track Err., q-1	-1.053 (0.15)	3.935 (0.65)	-3.481 (0.35)	24.473** (2.11)	30.043** (2.19)	-17.454 (1.08)	-13.767 (1.50)	-7.622 (1.20)	-0.386 (0.03)
ETF LIQ, q-1	10.253*** (12.20)	5.807*** (4.64)	4.382*** (2.91)	7.660*** (4.90)	1.521 (0.49)	6.113*** (3.12)	10.881*** (9.83)	7.396*** (5.69)	3.416* (1.74)
UND LIQ, q-1	1.458** (2.06)	2.058 (1.51)	-0.410 (0.31)	1.013 (1.11)	3.125** (2.01)	-2.240 (1.56)	1.052 (0.84)	-0.104 (0.05)	1.763 (0.79)
Dummies:									
Style, sector	YES	YES	YES	YES	YES	YES	YES	YES	YES
Calendar	YES	YES	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.430	0.304	0.202	0.342	0.319	0.264	0.404	0.365	0.212
N	3,608	3,594	3,592	1,667	1,665	1,664	1,941	1,929	1,928
<b>PANEL B: Investment companies and advisors (short-term) vs. bank trusts (long-term)</b>									
Exp., q-1	5.552 (0.98)	-3.284 (0.79)	8.096 (1.28)	9.465 (1.23)	17.234*** (3.04)	-8.036 (1.29)	-1.748 (0.22)	-18.124*** (3.47)	15.262* (1.70)
Age, q	1.096 (1.28)	0.916** (2.14)	0.129 (0.16)	1.100 (0.86)	-0.084 (0.14)	1.175 (1.13)	1.496 (1.47)	2.318*** (3.24)	-0.953 (0.76)
Age^2, q	-0.073 (1.30)	-0.042 (1.59)	-0.030 (0.54)	-0.073 (0.87)	0.009 (0.27)	-0.083 (1.29)	-0.071 (1.04)	-0.117** (2.46)	0.053 (0.60)
AUM, q-1	-2.503*** (3.68)	0.898* (1.87)	-3.406*** (4.45)	-0.987 (0.88)	2.686*** (3.99)	-3.665*** (3.70)	-3.790*** (4.23)	-0.616 (1.11)	-3.195*** (3.04)
ETF Ret, q-1	0.021 (0.62)	0.055** (2.02)	-0.032 (0.67)	0.093 (1.36)	-0.036 (0.84)	0.131** (2.10)	0.051 (1.37)	0.074** (2.59)	-0.025 (0.47)
Track Err., q-1	5.101 (0.70)	2.756 (0.86)	2.247 (0.27)	34.370*** (2.93)	5.683 (0.56)	28.759 (1.56)	-9.008 (1.03)	-0.376 (0.10)	-8.851 (0.87)
ETF LIQ, q-1	9.691*** (8.52)	2.903*** (2.97)	6.741*** (4.09)	6.898*** (2.66)	0.795 (0.55)	6.096*** (3.82)	10.160*** (7.47)	3.846*** (3.33)	6.257*** (2.88)
UND LIQ, q-1	2.704** (2.59)	0.473 (0.50)	2.175* (1.66)	0.920 (0.63)	2.212** (2.11)	-1.300 (0.81)	4.458** (2.25)	-1.436 (1.16)	5.757*** (2.84)
Dummies:									
Calendar	YES	YES	YES	YES	YES	YES	YES	YES	YES
Style, sector	YES	YES	YES	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.307	0.288	0.159	0.237	0.367	0.194	0.329	0.322	0.207
N	3,609	3,599	3,598	1,667	1,667	1,666	1,942	1,932	1,932

**Table 2- 11: Short-term trading and liquidity**

This table reports the results for a regression of average institutional investors' holding period (called ETF duration) on measures of liquidity (ETF, underlying portfolio and relative liquidity) and fund characteristics. ETF (underlying portfolio) liquidity is calculated as the first principal component of ETF (underlying portfolio) spreads, price impact and turnover. Relative liquidity is the first principal component of relative ETF liquidity (spreads, price impact, turnover and share creation activity). Inv. age is a proxy for the age of an ETF's investor base. All other variables have been defined in Table 9. \*\*\*/\*\*\*\* denotes statistical significance at the 10, 5 and 1 percent level. Standard errors are clustered by the fund (t-statistics in brackets).

<b>Variables</b>	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>
Inv. age, q-1	0.085*** (11.44)	0.086*** (11.36)	0.074*** (9.85)	0.076*** (9.86)	0.085*** (6.37)	0.086*** (6.43)
Exp. Ratio, q-1	-0.586* (1.76)	-0.611* (1.79)	-0.536* (1.72)	-0.538 (1.63)	1.490 (1.35)	1.493 (1.39)
Age, q	0.166*** (2.85)	0.165*** (2.86)	0.162** (2.53)	0.173*** (2.66)	0.382*** (5.35)	0.374*** (5.26)
Age^2, q	-0.015*** (3.77)	-0.015*** (3.78)	-0.016*** (3.53)	-0.017*** (3.61)	-0.025*** (6.43)	-0.025*** (6.71)
AUM, q-1	0.360*** (9.34)	0.366*** (9.65)	0.396*** (9.98)	0.389*** (9.95)	0.301*** (4.20)	0.307*** (4.36)
ETF Ret, q-1	-0.009*** (6.73)	-0.009*** (6.93)	-0.003 (1.28)	-0.002 (1.16)	-0.003 (1.38)	-0.002 (1.33)
Track Err., q-1	-0.673* (1.86)	-0.894*** (2.71)	-0.081 (0.22)	-0.142 (0.41)	0.642* (1.73)	0.421 (1.27)
ETF LIQ, q-1	-0.579*** (9.77)		-0.547*** (10.01)		-0.383*** (5.45)	
UND LIQ, q-1	0.201*** (2.62)		0.089 (1.15)		0.191 (1.61)	
REL LIQ, q-1		-0.611*** (10.38)		-0.556*** (10.10)		-0.434*** (5.75)
<b>Dummies:</b>						
Style, sector	YES	YES	YES	YES	YES	YES
Calendar	NO	NO	YES	YES	YES	YES
ETF	NO	NO	NO	NO	YES	YES
R <sup>2</sup>	0.571	0.568	0.607	0.602	0.724	0.725
N	3,278	3,278	3,278	3,278	3,278	3,278

## **CHAPTER THREE: LIQUIDITY, STYLE INVESTING AND EXCESS COMOVEMENT OF EXCHANGE-TRADED FUND RETURNS**

### **3.1 INTRODUCTION**

Liquidity is generally considered to be beneficial for pricing efficiency because it facilitates arbitrage. In this study I argue that liquidity can sometimes also be detrimental for pricing efficiency because liquidity facilitates short-term trading that has the potential to generate excess comovements among asset returns. In Barberis and Shleifer (2003) investors allocate money at the style level and engage in short-term style switching for reasons unrelated to fundamentals – allocating more capital to styles that recently performed well and taking money out of styles that have done poorly. This type of correlated trading can induce a common factor in the returns of assets in the same style<sup>20</sup>.

Investor demand should go first to the securities where the purest play exists and where liquidity is highest. Exchange-Traded Funds provide investors with easy access to popular investment styles (e.g. Large, Small, Value, Growth and Sector) at a cost that is on average lower

---

<sup>20</sup> Similar predictions arise in preferred habitat model of excess comovement (Barberis, Shleifer and Wurgler, 2005), which predicts that some investors restrict their trading to a subset of securities and the correlated non-fundamental demand of these investors is responsible for generating excess comovements.

relative to their underlying basket of securities (see Chapter 2). Moreover, it is easy to move money in and out of two different styles with ETFs and to enter into long-short strategies (e.g. Value-Growth) due to the relatively low short-selling costs of ETFs.

My conjecture is that, due to the ease of investing in investment styles with ETFs and because of their high liquidity, ETFs attract a clientele of short-term investors with correlated non-fundamental demand at the style level. Consequently, the returns of ETFs will be more exposed to a common source of style-based non-fundamental risk *relative* to their underlying securities. This relative, or twin-based, comparison allows me to identify excess comovements by studying common factors in the change in misvaluation, proxied by the return difference between an ETF and its underlying portfolio Net Asset Value (NAV). This approach is in sharp contrast to existing studies that investigate anomalous return comovements around “exogenous” events, or by relying on a CAPM type model to filter out the fundamental component of returns<sup>21</sup>. Moreover, by properly controlling for fundamental drivers of return comovements, I can examine what affects the degree of excess comovements in order to provide a better understanding of the ETF characteristics, particularly liquidity, that facilitate excess comovements.

An alternative mechanism that can generate excess comovements is differences in the speed of information diffusion between ETFs and their underlying portfolios. In this case the high liquidity of ETFs is more likely to attract investors with fundamental (long-term) information about abstract risk factors. Differences in information diffusion can also arise mechanically when there is stale pricing in the underlying securities (e.g. in small-cap stocks). This hypothesis is also known as the information diffusion view of excess comovements (see Barberis, Shleifer and Wurgler, 2005).

An important distinction between the non-fundamentals-based and the fundamentals-based view of excess comovement is that the former assumes that style investor have short horizons. Although the high liquidity of ETFs is beneficial to both long- and short-term investors, I argue that liquidity benefits short-term investors the most as in Amihud and Mendelson (1986). Supporting this conjecture, in Chapter 2 I show that ETFs with high liquidity relative to their underlying securities have higher fund flows in the short-term, higher institutional ownership by

---

<sup>21</sup> e.g. Barberis, Shleifer and Wurgler (2005), Prinsky and Wang (2006), Green and Hwang (2009), Kumar, Page and Spalt (2013)

short-term (relative to long-term) investors and shorter institutional holding periods (relative to their underlying baskets). Retail investors are even more likely to be attracted to ETFs for liquidity reasons because the transaction costs that they face when investing in the underlying security basket are likely prohibitive.

To make my tests as clean as possible, I focus on a sample 164 physically replicated ETFs that are traded in the U.S. and that track only U.S. equity indices. These funds have over \$540 billion in total assets as of 12/2012 – roughly 85 percent of the total assets of all U.S. equity ETFs. In contrast to related studies on “twin securities”; cross-listed stocks (e.g. Gagnon and Karolyi, 2010), international closed-end funds (Bodurtha, Kim and Lee, 1995), or even domestic closed-end funds (Lee, Shleifer and Thaler, 1991), my sample is unlikely to be affected by either non-synchronicity or stale pricing. The former is not a concern since ETFs and their underlying securities are traded in the same time-zone. Stale pricing is unlikely to occur because both ETFs and their underlying securities are generally actively traded, with the possible exception of small-cap stocks. I conduct several tests based on reversals in misvaluation to rule out this possibility.

To preview my results, I find significant evidence of commonality in misvaluation at the investment style level (size, valuation and sector): changes in misvaluation (ETF-NAV returns) comove positively across ETFs in similar styles, and negatively with ETFs in distant styles, after controlling for arbitrage induced mean-reversion in misvaluation. To illustrate the economic magnitude, a one Std. Dev. increase in the own-style misvaluation factor is on average associated with an increase in daily ETF-NAV return differentials of 55.73 percent of the Std. Dev. of ETF-NAV returns. The impact of a one Std. Dev. shock to the own-style factor is also considerable relative to the variability in raw returns at roughly 4 percent<sup>22</sup>, but declines with the return horizon to 2.29 and 1.27 percent in weekly and monthly data respectively. Despite the decline in magnitudes over time, the results remain highly significant even in monthly data, which is more consistent with the non-fundamentals based view of excess comovement as opposed to information diffusion, because the latter predicts that information is assimilated rapidly to both ETFs and their underlying securities since both are liquid and actively traded instruments.

The evidence of style-based commonality in misvaluation implies excess comovement of returns, but of which, the ETF or the NAV? To provide more direct evidence of misvaluation in

---

<sup>22</sup> Calculated as  $\beta_{OWN,i} * Std(Own\ style\ factor) / Std(Raw\ return\ of\ ETF\ i)$ , averaged across all ETFs.

the ETF leg, rather than the NAV, I investigate return reversals. Specifically, if an ETF is hit by a positive *non-fundamental* demand shock that pushes its price above the NAV value (resulting in a positive ETF premium), then we should observe a reversal in the future ETF returns (as a result of arbitrage) without a meaningful impact on NAV returns. Conversely, if the initial positive demand shock was driven by positive *fundamental* news that is absorbed first into ETF prices, then future NAV returns should be positive as the NAV catches up with a lag and the ETF return should remain unaffected. The empirical results confirm that ETF premiums have a negative and significant impact on future ETF returns over a period of three to four days, consistent with premiums reflecting non-fundamental demand shocks. More importantly, reversals in ETF returns are strongest among small-cap ETFs (the category with the highest liquidity relative to their underlying basket), which is consistent with the conjecture that liquidity attract short-term noise traders with correlated demand for investment styles.

Moreover, current ETF premiums also forecast future NAV returns negatively over a four day period (positive on first day, negative on the remaining days), which is opposite to what the information diffusion view would predict. Such a negative relationship can, however, arise when investors experience non-fundamental demand shocks and trade sequentially. In this case liquidity goes first to the most liquid securities (ETFs) and when liquidity dries up, demand goes to the next most liquid ETF and so on, until no more ETFs are sufficiently liquid relative to their underlying securities, in which case the demand goes to the underlying securities.

To provide further evidence that commonality in misvaluation is driven by commonality in demand shocks, I begin by investigating commonality in turnover and liquidity, which have been linked to correlated trading (e.g. Chordia, Roll and Subrahmanyam, 2000; Karolyi, Lee and Van Dijk, 2012). I find similar style-based comovements in relative measures for turnover and liquidity (i.e. ETF minus underlying portfolio) with most of the effect originating from ETF leg, rather than the NAV. Next, I establish that commonality in demand shocks can predict one-month ahead commonality in misvaluation, which is consistent with the idea that excess comovements are driven by correlated non-fundamental demand shocks. Finally, I investigate the determinants of the degree of commonality in misvaluation and find that ETFs with more desirable liquidity characteristics have a greater degree of commonality in misvaluation. This is to be expected if liquidity is what attracts short-term traders to ETFs.

Understanding what affects asset prices in the ETF market is important due to the potential for spillovers across markets. Staer (2014) shows that ETF fund flows have a large impact on underlying stock returns, almost half of which is reversed within a few days. Ben-David, Franzoni and Moussawi (2014) find that higher ETF ownership of stocks is associated with more volatile stock returns and a stronger mean-reverting component in stock returns, while Da and Shive (2013) link higher ETF ownership to stronger underlying stock return comovements. My conjecture that ETFs attract high-turnover investors with correlated trading needs is consistent with these findings.

Among the most widely cited evidence in favor of correlated demand-based theories of excess comovement are the comovements observed around index additions (with other index stocks) and stock splits (with low-priced stocks)<sup>23</sup>. The critical assumption, that the event is exogenous, remains controversial and has recently been challenged by Kasch and Sarkar (2012) and Perez, Shkilko and Tang (2012). A broader debate in the literature concerns whether the observed comovement patterns among small-cap stocks (Banz, 1981) or value/growth stocks (Fama and French, 1993, 1995) can be explained by common variation in cash flows or discount rates<sup>24</sup>; or by unmodeled irrational behavior (see Barberis and Thaler, 2003), and to what extent limits-to-arbitrage can explain these findings (Brav, Heaton, Li, 2010). My contribution in this regard is to provide a more controlled experiment that is better suited for separating the two sources (fundamental vs. non-fundamental) of return comovements.

This paper is also related to a growing literature on the relationship between correlated trading and return comovements. Kumar and Lee (2006) find not only that retail trades are systematically correlated, but also that such trades can help explain some of the anomalous return comovements among stocks with high arbitrage costs. Correlated retail demand has also been linked to investors' tendency to place similar speculative bets (Dorn, Huberman and Sengmueller, 2008). Kumar, Page and Spalt (2013b) show that stocks with lottery-like feature comove too much with one another due to the correlated trading activity of gambling-motivated investors. Greenwood (2007) constructs a simple trading strategy that bets on the reversion of the prices of over-weighted Nikkei 225 stocks that comove too much in the short-run and finds this trading strategy to yield significant risk-adjusted profits.

---

<sup>23</sup> See e.g. Barberis, Shleifer and Wurgler (2005), Green and Hwang (2009), Kumar, Page and Spalt (2013).

<sup>24</sup> See e.g. Fama and French (1993), (1995); Campbell, Polk and Vuolteenaho (2009); Campbell *et al.* (2013)

The article proceeds as follows. Section 2 provides some background information on ETF arbitrage and institutional details. Section 3 provides the theoretical framework and presents the main testable implications. Section 4 describes the data, defines the key variables and presents summary statistics. Section 5 presents the empirical tests for excess comovement based on an analysis of commonality in ETF misvaluation. Section 6 establishes that ETF premiums mainly reflect misvaluation in the ETF, rather than NAV leg. Section 7 documents that measures of ETF demand shocks also exhibit style-based comovement and that the degree of common demand shocks can predict commonality in misvaluation, along with ETF characteristics associated with higher liquidity. Section 8 concludes.

### **3.2 BACKGROUND ON ETF ARBITRAGE AND INSTITUTIONAL DETAILS**

ETFs have an open-ended structure via the share creation and redemption process that facilitates arbitrage. This process is only available to some institutional investors (called Authorized Participants, or APs), which have signed an agreement with the ETF sponsor. APs can buy or sell ETF shares in bundles (or creation units) directly from the ETF sponsor in exchange for the underlying basket of securities at the end of the trading day (at 4 P.M. EST). Although this process is limited to APs (typically market makers, broker/dealers or large institutions), they can also create (or redeem) shares directly for their clients who wish to transact in ETFs.

To illustrate the arbitrage process via the share creation mechanism, consider a situation where the ETF is trading at a premium (ETF price is above the NAV). An AP would then buy the underlying basket (at the NAV), exchange the basket for new ETF shares with the ETF sponsor and sell the newly created shares on the secondary market. In practice an AP could buy the underlying basket and simultaneously short the ETF. At the end of the trading day the AP would exchange the underlying basket for ETF shares, and close the short ETF position.

The direct costs of creating ETF shares are small for U.S. equity funds (the focus of this paper). The size of a creation unit is typically 50,000 or 100,000 shares with dollar values ranging from \$300,000 to \$10 million. The fixed creation costs range from \$500 to \$3,000. For SPY, the world's largest and most actively traded ETF tracking the S&P 500, the fixed fee of \$3,000 amounts to about 5 bp for one creation unit worth \$6 million, or 1 bp for five creation units worth about \$30 million (Petajisto, 2013). For a sample of equity U.S. ETFs<sup>25</sup>, in Chapter 2

---

<sup>25</sup> Their sample is identical to mine. More details appear in the data section.

I report that share creations/redemptions occur on 30.9 (22.7) % of trading days on average (median) and conditional on such days, the magnitudes are \$69.6 million (\$12.4 million) or 244.3 percent (27.4 percent) of daily dollar volume. These magnitudes indicate that AP's frequently create/redeem multiple creation units at a given point in time, possibly to reduce costs.

Arbitrage activity is also undertaken by market participants other than APs, such as hedge funds and high-frequency traders (Marshall, Nguyen, and Visaltanachoti, 2013). For instance, when the ETF is trading at a premium, an investor can purchase the underpriced asset (NAV), short-sell the overpriced asset (ETF) and wait for prices to converge to realize an arbitrage profit. ETF prices can also be arbitrated against other ETFs (Marshall, Nguyen, and Visaltanachoti, 2013; Petajisto, 2013) or against futures contracts (Richie, Daigler, and Gleason, 2008).

### **3.2.1 Limits-to-arbitrage**

Despite the enhanced pricing efficiency of ETFs, arbitrage remains limited by transactions costs, holding costs and other implicit restrictions (e.g. minimum share creation). As for transactions costs, both ETF and underlying portfolio spreads matter because arbitrage trades require access to both markets. Price impact is also of particular concern, as shown by Staer (2014). The potentially high price impact costs of ETF share creations combined with the large size of typical creation events (see Chapter 2) indicate that APs might need several days to accumulate a position that is big enough to offset the creation without undue price impact. This makes it harder to trade on small price deviations. Traditional long-short arbitrage trades with smaller trade sizes can be used to avoid some of the price impact costs. However, such arbitrage trades are exposed to holding costs (costs that accrue every period a position), especially idiosyncratic risk, for as long as the arbitrage trade is kept open (see Pontiff, 2006).

Greenwood's (2005) model can be used to justify limits-to-arbitrage further. In their model market-makers (or APs in the ETF market) are risk-averse and require compensation for providing liquidity. Similar predictions arise in Cespa and Foucault's (2014) model with multiple investor classes and some degree of market fragmentation. Ben-David, Franzoni and Moussawi (2014) discuss a dynamic extension of Cespa and Foucault's (2012) model to further justify temporary price discrepancies between identical assets.

### 3.3 THEORETICAL FRAMEWORK AND TESTABLE IMPLICATIONS

The theoretical channel for excess comovement in ETF returns relies on limited arbitrage, correlated demand and clientele effects. In the Barberis and Shleifer (2003) model, investors allocate funds at the style level (e.g. small or value) as opposed to at the individual asset level. The strong demand for investment styles is evident from the large number of ETFs, mutual funds, and hedge funds that follow distinct styles and which are used by both individual and institutional investors<sup>26</sup>. The style investing model also assumes that investors move into styles that have performed well in the past, and out of styles that have performed poorly. If some of these style investors are also noise traders with correlated sentiment (Baker and Wurgler, 2006), then coordinated shifts in investor preferences across investment styles (e.g. from value to growth) will induce a common factor in the returns of assets in the same style. In this case, the return of security  $i$  and  $j$ , belonging to style  $X$  and  $Y$  respectively, are given by<sup>27</sup>:

$$\begin{aligned} R_{i,t} &= \varepsilon_{i,t} + \gamma_{i,X} \Delta u_{X,t} + \gamma_{i,Y} \Delta u_{Y,t} \quad \text{where } i \in X \\ R_{j,t} &= \varepsilon_{j,t} + \gamma_{j,X} \Delta u_{X,t} + \gamma_{j,Y} \Delta u_{Y,t} \quad \text{where } j \in Y \end{aligned} \quad (1)$$

The first component reflects news about fundamental value ( $\varepsilon_t$ ), which is often characterized via an asset pricing model such as the CAPM or the intertemporal CAPM (Merton, 1973).  $\Delta u_{X,t}$  and  $\Delta u_{Y,t}$  reflect common *demand* shocks (or noise-trader sentiment) for securities in styles  $X$  and  $Y$  respectively. In this theoretical setup each security  $i$  is positively exposed to demand shocks in its own style ( $\gamma_{i,X} > 0$ ), and negatively exposed to demand shocks in distant styles ( $\gamma_{i,Y} < 0$ ). Another interpretation of Eq. (1) is that some investors restrict themselves to trading ETFs in a specific style giving rise to preferred habitats (Barberis, Shleifer and Wurgler, 2005). In this case  $\Delta u_t$  captures changes in sentiment, risk-aversion or liquidity needs of the style investors in habitat  $X$  or  $Y$ . As in Greenwood (2007), I refer to both interpretations as the non-fundamentals based view of excess comovement.

Investor demand should go first to the securities where the purest play exists and where liquidity is highest. Exchange-Traded Funds provide investors with easy access to popular investment styles at a cost that is on average lower relative to their underlying basket of securities (see Chapter 2). Moreover, it is easy to move money in and out of two different styles

---

<sup>26</sup> see e.g. Brown and Goetzmann (1997); Fung and Hsieh (1997); and Chan, Chen, and Lakonishok (2002)

<sup>27</sup> see Eq. (20) in Barberis, Shleifer and Wurgler (2002)

with ETFs and to enter into long-short strategies (e.g. Value-Growth) due to the relatively low short-selling costs of ETFs<sup>28</sup>.

My conjecture is that, due to the ease of investing in investment styles with ETFs and because of their high liquidity (see Chapter 2), ETFs attract a clientele of short-term investors with correlated non-fundamental demand for investment styles. Hence, the returns of ETFs in similar styles will comove excessively – i.e. after controlling for the return on the fundamental underlying portfolio Net Asset Value (NAV) – with one another. To arrive at a testable hypothesis, I first take the return difference between ETF  $i$  and its NAV:

$$R_{i,t}^{ETF} - R_{i,t}^{NAV} = (\varepsilon_{i,t}^{ETF} - \varepsilon_{i,t}^{NAV}) + (\gamma_{i,X}^{ETF} - \gamma_{i,X}^{NAV}) \Delta u_{X,t} + (\gamma_{i,Y}^{ETF} - \gamma_{i,Y}^{NAV}) \Delta u_{Y,t} \quad (2)$$

where  $R_i$  = return for ETF  $i$  or its underlying portfolio NAV at time  $t$

$\gamma_i$  = sensitivity to demand shocks of ETF  $i$  or its underlying portfolio NAV

The return difference (2) can be interpreted as proxy for the change in misvaluation (from an arbitrageurs point of view) because an ETF is a claim on its underlying portfolio of securities. Both securities should therefore be equally exposed to fundamental news ( $\varepsilon_{i,t}^{ETF} - \varepsilon_{i,t}^{NAV} = 0$ ). This does, however, not rule out the possibility that both securities are misvalued at the same point in time. In section 5.3, I rule out the possibility that Eq. (2) is driven by differences in systematic risk between the ETF and the NAV. The testable implication of style-based *excess comovement* is as follows:

**Hypothesis 1:** *Changes in misvaluation (Eq. (2)) of two ETFs in the same style ( $i, j \in X$ ) are positively correlated,  $\text{corr}(R_i^{ETF} - R_i^{NAV}, R_j^{ETF} - R_j^{NAV}) > 0$ , while changes in misvaluation of two ETFs in distant styles ( $i \in X, k \in Y$ ) are negatively correlated,  $\text{corr}(R_i^{ETF} - R_i^{NAV}, R_k^{ETF} - R_k^{NAV}) < 0$ .*

To illustrate, investors increase their allocation to a particular style (say Value) after it has outperformed its distant style (Growth). This switch in allocations is financed either by selling securities in the distant style (Growth), or by selling securities in every other style (everything except Value). These demand shocks can then generate a positive correlation between changes in misvaluation of Value ETFs, and a negative correlation between Value and Growth ETFs.

---

<sup>28</sup> “No Shortage of Share Lending” featured in Journal of Indexes, February 17, 2010.

Style-based excess comovement can also arise if there are differences in the speed of information diffusion between the ETF and the NAV (Barberis, Shleifer and Wurgler, 2005). Specifically, because ETFs are highly liquid and represent pure plays on investment styles, fundamental news about abstract risk-factors may be incorporated first into ETF prices, and with a lag to NAV prices. In this case the ETF and NAV returns are given by:

$$\begin{aligned} R_{i,t}^{ETF} &= \varepsilon_{i,t} \\ R_{i,t}^{NAV} &= \mu\varepsilon_{i,t} + (1-\mu)\varepsilon_{i,t-1} \end{aligned} \tag{3}$$

where  $\mu$  = fraction of news incorporated with a lag

In contrast to the information diffusion view, the argument that ETFs attract short-term investors with correlated non-fundamental demand relies on liquidity clienteles, formalized by Amihud and Mendelson (1986). Their model predicts that short-horizon investors self-select into more liquid assets, such as ETFs. Supporting this conjecture, in Chapter 2 I show that higher ETF liquidity predicts fund flows strongly over short horizons (weekly and monthly), while expense ratios matter the most over longer horizons. Amongst institutional investors, they also show that funds with higher liquidity experience increased ownership by short-term investors relative to long-term, more institutional buying and more selling over the following quarter, and shorter holding periods.

As for retail investors, the argument for liquidity clienteles is even stronger because the transactions costs that they face when investing in the underlying security basket are likely prohibitive in comparison to ETFs. Moreover, ETFs generally have lower expense ratios than even their cheapest retail mutual fund counterparts. Retail investors do pay attention to salient trading costs such as front-end loads and commissions (Barber, Odean and Zheng, 2005) as well as expense ratios (Grinblatt *et al.*, 2013) in the case of mutual funds. For ETFs, the most salient costs are likely to be quoted spreads and expense ratios, both of which are widely disseminated, while commissions are generally small, and sometimes even close to zero<sup>29</sup>.

One way to separate the causes of excess comovements is to investigate its degree of persistence. According to the information diffusion view, excess comovements are unlikely to persist for long (e.g. over a week or a month) because both ETFs and their underlying securities are liquid and should therefore incorporate news relatively fast (e.g., the delay parameter in (3)

---

<sup>29</sup> Many ETFs have free commissions: for a list see <http://etfdb.com/type/commission-free/all/>.

could be a day or less for S&P 500 stocks). In contrast, the non-fundamental based view suggests that excess comovement may persist over longer horizons (e.g. monthly or quarterly) because styles go through cycles (Barberis and Shleifer, 2003). Empirical evidence also suggests that investors allocate funds based on past relative style performance evaluated over monthly and quarterly periods (Broman and Shum, 2015b).

In order to properly disentangle the two stories we can directly examine the source of misvaluation (ETF vs. NAV). The non-fundamentals based view predicts that ETFs are hit by temporary demand shocks that subsequently revert. To illustrate the mechanism, suppose that noise-traders buy value ETFs due to their high recent performance and liquidity. The resulting price pressure is associated with an increase in the ETF price above its fundamental value (NAV). Over time arbitrageurs trade on the misvaluation and push the ETF price back down to its fundamental value.

**Hypothesis 2a:** *Current ETF premiums (ETF-NAV price deviations) predict future ETF returns negatively, while future NAV returns remain unaffected.*

In contrast, the information diffusion view predicts that ETFs impound fundamental information first, while their underlying securities (NAV) catch up with a lag. In this case investors with fundamental information about abstract style-based risk-factors trade first in ETFs due to their high liquidity and pure play resulting in positive price pressure. Subsequently, the new fundamental information will also be incorporated in the prices of the underlying securities, resulting in positive price pressure on the NAV.

**Hypothesis 2b:** *Current ETF premiums (ETF-NAV price deviations) predict future NAV returns positively, while future ETF returns are not affected.*

Differences in information diffusion can also arise mechanically when there is stale pricing in the underlying securities (NAV is based on closing prices which may be stale). Stale pricing could be a concern for some small-cap stocks. To rule out this possibility, I examine whether the results for commonality in misvaluation continue to hold among ETFs that hold underlying securities that are not prone to stale pricing (e.g. large-cap).

Finally, if the degree of commonality in non-fundamental demand shocks varies across ETFs or over time, then this variation should be positively related to the degree of excess comovements (Greenwood, 2007; Greenwood and Thesmar, 2011). Moreover, with an

appropriate adjustment for total misvaluation, excess comovements of ETF  $i$  should also be stronger for more liquid ETFs because they are more likely to attract short-term noise traders.

**Hypothesis 3:** *The degree of commonality in misvaluation is positively related to the degree of commonality in non-fundamental demand and the liquidity characteristics of ETF  $i$ .*

### 3.3.1 Key variables

ETF misvaluation can be measured by the premium, or the log-difference between the market price of an ETF and the market value of the ETF's portfolio on a per-share basis (NAV<sup>30</sup>):

$$PREM_{i,t} = \ln(P_{i,t}^{ETF}) - \ln(P_{i,t}^{NAV}) \quad (4)$$

where:  $P_{i,t}^{ETF}$  = bid-ask midpoint price for ETF  $i$  at the end of day  $t$ ,

$P_{i,t}^{NAV}$  = underlying portfolio Net Asset Value per share for ETF  $i$  on day  $t$

Hypothesis 1 predicts that changes in misvaluation ( $R_{i,t}^{E-N}$ ) are correlated across ETFs in similar styles. When log-returns are used,  $R_{i,t}^{E-N}$  corresponds to the change in premium:

$$\Delta PREM_{i,t} \equiv \ln(P_{i,t}^{ETF} / P_{i,t}^{NAV}) - \ln(P_{i,t-1}^{ETF} / P_{i,t-1}^{NAV}) = R_{i,t}^{ETF} - R_{i,t}^{NAV} \quad (5)$$

I will use log-returns throughout this study to keep the link clear between ETF-NAV return differences and changes in premiums.

## 3.4 DATA

My data selection starts with all U.S. traded Exchange-Traded Funds that exist both in Bloomberg and in Morningstar Direct. I keep funds that i) invest in U.S. equity, ii) are physically replicated and “passively” managed and iii) have at least 3 years of data available<sup>31</sup>. The second criterion, which excludes actively managed and synthetically replicated ETFs, is particularly important for this study. Active ETFs include ETFs designated by Morningstar as following active strategies, “smart beta” funds that give investor's access fundamentally-weighted indices and funds based on proprietary underlying indices. These are excluded because they are less likely to represent pure plays on investment styles, they may be chosen by investors because of their investment strategy or manager performance and also because their holdings change

<sup>30</sup> NAV also includes accrued income from securities lending, underlying stock dividends and cash.

<sup>31</sup> I exclude the first 6 months of a funds history since the data can be unreliable (see Chapter 2), leaving me with an estimation sample of at least 2.5 years.

frequently making it more difficult to measure their NAVs and misvaluation. Synthetically replicated ETFs (i.e. leveraged, inverse and futures-based ETFs) are excluded because their holdings change frequently, and because arbitrage is more risky as share creations are settled in cash<sup>32</sup> rather than in-kind. These three exclusion criteria decrease the sample of ETFs from 363 to 224 to 164. In terms of assets under management (AUM), the total AUM of U.S. ETFs was \$632 billion in September 2012 according Blackrock (2012), while the AUM of my 164 funds was \$540 billion. Given the dramatic expansion in scope and size of the ETF market in the last five to ten years, earlier data may not be as representative of current market conditions, which is why I decided to focus on a recent sample period, from January 2006 to December 2012. I also conduct robustness tests on a sample starting in June 2002.

The sample of ETFs, along with NAVs<sup>33</sup>, shares outstanding and underlying indices, is obtained from Bloomberg. To ensure the reliability of NAV prices from Bloomberg, I cross-check the NAV data with CRSP Mutual Fund Database. Specifically, NAV prices from Bloomberg sometimes represent stale values from the previous trading day, in which case I use the NAV price obtained from the CRSP Mutual Fund Database. My second source is CRSP, which I use to obtain price, return and volume data for all funds.

The ETF and NAV price data contains a handful of extreme observations and other outliers that need to be dealt with. ETF premiums greater than 20 % are mainly due to data errors (Petajisto, 2013), and are therefore discarded. Moreover, when the premium based on midpoint prices is more than ten percentage points greater in absolute terms than the premium based on closing prices, I use the latter instead. Finally, levels and changes in premiums are winsorized fund-by-fund at 5 Std. Dev. from the mean to reduce the impact of any remaining outliers.

Third, I use the ubiquitous 3-by-3 Morningstar style classification (Small-, Mid- and Large; Value-Blend-Growth) to identify the investment style of a fund. I use the size and valuation styles based on the evidence in Froot and Teo (2009) and Kumar (2009) that both retail and institutional investors allocate capital at the size and value-growth level. The Morningstar classification has three key advantages. First, it coincides with the dichotomy often used by practitioners and many ETFs are named after their Morningstar style analogs (e.g. SPDR S&P

---

<sup>32</sup> With cash settlement arbitrageurs are exposed to idiosyncratic risk because many ETFs have a cut-off time in the afternoon to submit creation orders implying that arbitrageurs do not get to see the end-of-day NAVs before making the decision to trade.

<sup>33</sup> For ETF's by the iShares provider I use the NAV data that is directly available from their website as they contain fewer data errors, as suggested by Petajisto (2013).

600 Small-Cap Value or iShares Russell 3000 Growth fund). Investors do pay attention to fund names as illustrated by Cooper, Gulen and Rau (2005). They show that mutual funds that take rename their fund to match the current “hot style “subsequently experience abnormal inflows, even when the name change is unrelated to performance or any real change in holdings to match the new style. Second, Morningstar is a leading fund information provider and its classification system is publicly available. Third, the Morningstar style classification is updated monthly and it is based entirely on firm characteristics, which yields more stable style classifications over time as opposed to a latent variable approach based on the Fama and French (1993) factor loadings.

Table 3-1 gives snapshots of the sample used in this study. At the beginning (01/2006), my sample contains 95 ETFs with \$199.82 billion in AUM. Subsequently there are 156 ETFs with \$257.55 billion in AUM (06/2009), and 164 ETFs with \$540.01 billion at the end of the sample (12/2012). Roughly half of the funds are in diversified non-sector styles, while the rest are sector ETFs. Among the non-sector ETFs and within each size-category, the number of blend funds (neither value, nor growth) is roughly equal to the number of value and growth funds combined. Sector ETFs are generally much smaller and account overall for only one third of the total AUM. According to Bloomberg, 142 ETFs are fully replicated (i.e. hold more than 90 percent of the securities in the underlying index), while the remaining 22 use physical sample replication.

[Table 3-1]

### 3.4.1 Descriptive statistics

Table 3-2 provides descriptive statistics for ETF premiums and ETF-NAV returns (changes in premiums). Both are zero on average, and at the median, which suggests that ETFs are overall efficiently priced. There is, however, considerable variation around the mean as indicated by the standard deviation of 0.09 % and 0.12 % for the level and change in premiums. The extreme right and left tails (1 and 99 percentiles) are roughly +/- 27 bps for levels of premiums, and +/- 37 % for changes in premiums. Another way to illustrate the magnitude of misvaluation is to calculate the variability in changes in misvaluation relative to the variability in raw ETF returns, or  $SD(R_{i,t}^{E-N})/SD(R_{i,t}^{ETF})$ . This equals a considerable 7.4 percent on an equally-weighted basis.

These numbers are based on mid-point ETF prices at the end of the day. In contrast, the “actual” premiums based on closing prices are almost twice as volatile indicating that the true

cost of trading against ETF misvaluation can be much higher. I use mid-point prices in order to mitigate concerns about the illiquidity of the shares of smaller ETFs (Engle and Sarkar, 2006).

[Table 3-2]

### 3.5 EMPIRICAL TESTS OF EXCESS COMOVEMENT

Style-based excess comovement of ETF returns (Hypothesis 1) predicts a positive correlation between the change in misvaluation of ETFs in similar styles ("OWN style"), and a negative correlation with other ETFs in distant styles ("DIST style"). This hypothesis can be conveniently tested by investigating common factors in misvaluation (or commonality in misvaluation):

$$R_{i,t}^{E-N} = \alpha_i + \lambda_i PREM_{i,t-1} + \beta_{i,O} R_{OWN,t}^{E-N} + \beta_{i,DI} R_{DIST,t}^{E-N} + e_{i,t} \quad (6)$$

where:  $R_{i,t}^{E-N}$  = change in ETF misvaluation, or  $R_{i,t}^{ETF} - R_{i,t}^{NAV}$ ,  $i \in OWN$

$R_{OWN,t}^{E-N}$  = own-style misvaluation factor,  $\sum_{j=1}^J w_{j,t} (R_{j,t}^{ETF} - R_{j,t}^{NAV})$ ,  $j \in OWN, j \neq i$

$R_{DIST,t}^{E-N}$  = distant-style misvaluation factor,  $\sum_{l=1}^L w_{l,t} (R_{l,t}^{ETF} - R_{l,t}^{NAV})$ ,  $l \in DIST$

$w_{j,t}$  = weight for ETF  $j$  at time  $t$

Hypothesis 1 predicts a positive correlation between the change in misvaluation of ETF  $i$  ( $R_{i,t}^{E-N}$ ) and the own-style misvaluation factor ( $R_{OWN,t}^{E-N}$ ,  $\beta_{i,O} > 0$ ). Note that the factor excludes ETF  $i$  to avoid inducing a spurious correlation. As discussed in the previous section, some style dimensions only have a few funds (e.g. 3 ETFs with a Small-Value classification). In order to obtain a parsimonious metric for the own-style factor, and to avoid producing noisy factors based on a few funds, I use the following weights ( $w_{j,t}$ ) in constructing the own-style factor: equal weight is given to funds that match both style dimensions (size and valuation), and half the equal weight to funds that are in adjacent styles. For instance, if ETF  $i$  is Large-Value, then adjacent styles include Mid-Value and Large-Blend. Blend funds (neither value, nor growth) are matched only by their size category. This approach is used for all diversified non-sector ETFs.

The Morningstar 3-by-3 style classification only considers size and valuation, while sector (or industry) styles may also be important. Froot and Teo (2008) show that institutional investors reallocate capital across three style dimensions (size, valuation and industry), while Choi and Sias (2009) provide evidence of industry herding among institutional investors that is distinct from, and at least equally important to, herding by size and valuation styles. For sector ETFs it may therefore be important to account for all three style dimensions. In constructing the own-

style factor for sector ETFs, I therefore give equal weight to other funds in the same industry, half the equal weight to any fund in the same size and valuation style and one fourth of the equal weight to any fund in adjacent Morningstar 3-by-3 styles. For the matching by Morningstar 3-by-3 style, I do not differentiate between sector and non-sector ETFs.

Hypothesis 1 also predicts a negative correlation between the change in misvaluation of ETF  $i$  ( $R_{i,t}^{E-N}$ ) and the distant-style misvaluation factor ( $R_{DI,t}^{E-N}, \beta_{i,DI} < 0$ ). As with the own-style factor, the distant-style factor is specific to ETF  $i$  and it is based on the weighted average ETF-NAV return of other ETFs in distant styles (relative to  $i$ ). The following weights ( $w_{l,t}$ ) are used: equal weight is given to funds that are in distant styles (along both the size and valuation dimensions) and half the equal weight is given to funds that are in styles adjacent to the distant style. For instance, if ETF  $i$  is Large-Value, then Small-Growth is the distant style (equal-weight) and Mid-Growth and Small-Blend (half the equal weight) are in styles adjacent to Small-Growth. For mid-cap funds I consider large and small to be equally distant. Hence, the distant style of Mid-Value is Small-Growth and Large-Growth<sup>34</sup>.

Even if the ETF-NAV return differentials perfectly control for fundamental risk (as in Eq. (2)), then the return difference reflects not only the impact of noise-trader demand shocks, but also the subsequent arbitrage trades and the associated reversal in misvaluation. Suppose that ETF  $i$  becomes overpriced relative to the NAV ( $PREM_{i,t} > 0$ ) as a result of a positive noise-trader demand shock. In this case arbitrageurs will buy the underpriced basket of securities, exchange it for ETF shares and sell the newly created ETFs in the secondary market, which will induce a reversal in misvaluation ( $\Delta PREM_{i,t+1} < 0$ ). This mechanism generates mean-reversion in changes in misvaluation. Given that demand shocks are theoretically expected to be uncorrelated (or positively auto-correlated due to some persistence in demand shocks), we can isolate the impact of demand shocks on misvaluation by controlling for the lagged premium in regression (6).

I run time-series regressions of (6) separately for each ETF using all available observations and report the mean of the estimated coefficients across all ETFs. In calculating the standard error for the mean coefficient, I take into account the cross-equation correlations in the estimated coefficients following Hameed, Kang and Viswanathan (2010):

---

<sup>34</sup> The results also hold if I construct own- and distant-style factors by matching on both the size and valuation styles while disregarding adjacent styles, or if I disregard the industry styles and match all ETFs (whether sector or non-sector) based on their size and valuation characteristics.

$$Std.Dev.(\bar{\beta}) = Std.Dev.\left(\frac{1}{N} \sum_{i=1}^N \beta_i\right) = \frac{1}{N} \sqrt{\sum_{i=1}^N Var(\beta_i) + \sum_{i=1}^N \sum_{j=1, j \neq i}^N \rho_{i,j} \sqrt{Var(\beta_i)Var(\beta_j)}} \quad (7)$$

where  $\sqrt{Var(\beta_i)}$  = the White standard error of the coefficient

$\rho_{i,j}$  = the estimated correlation between the residuals for ETF  $i$  and  $j$ .

### 3.5.1 Results: style-based commonality in misvaluation

The main results for style-based commonality in misvaluation (regression (6)) are given in Table 3-3. The results are reported separately for the daily, weekly and monthly horizons. For weekly and monthly data, I require at least 75 percent valid daily observations in each period.

[Table 3-3]

The results in Table 3-3 for the daily horizon show not only that the own-style betas are on average positive and significant, but also that more than 95 percent of the betas are positive and (individually) significant. The economic magnitudes are also considerable. To illustrate, a one Std. Dev. increase in the own-style misvaluation factor is associated with a 6.99 bps increase in ETF-NAV returns. The proportional impact,  $\beta_{OWN} * STD(R_{OWN,t}^{E-N}) / STD(R_{i,t}^{E-N})$  is also considerable at 55.73 percent on average indicating that (loosely speaking) almost half of the variation in changes in misvaluation is driven by the own-style misvaluation factor. The economic magnitudes, in terms of the proportion of common variation in misvaluation, remains very similar at weekly and monthly horizons. This finding is more consistent with the non-fundamentals based view of excess comovement as opposed to information diffusion (or stale pricing), because the latter predicts that information is assimilated relatively fast to both ETFs and their underlying securities since both are liquid instruments.

Since the price pressure associated with non-fundamental demand shocks is temporary, the strength of excess comovements should, on the one hand, decline with the length of the return horizon. On the other hand, if there is some persistence in demand shocks combined with limits-to-arbitrage, excess comovements may persist over longer horizons. To better assess whether commonality in misvaluation remains economically important over longer horizons, I recompute the proportional impacts as  $\beta_{OWN} * STD(R_{OWN,t}^{E-N}) / STD(R_{i,t}^{ETF})$ . This metric tells us how important common misvaluation is relative to the variability in raw returns. In daily data, common misvaluation accounts for roughly 4 percent of the variability in raw returns. This effect declines

to 2.29 percent in weekly data, and to 1.27 percent in monthly. Thus, while there is some persistence in the degree of commonality in misvaluation even over monthly horizons, the importance of common misvaluation declines over longer horizons consistent with arbitrage forces playing a role.

The prediction that changes in misvaluation are negatively correlated between ETFs in distant styles is also born out in the data as the distant-style betas are significantly negative ( $\beta_{DI} < 0$ ), although the average  $\beta_{DI}$  is only one tenth as large as the average  $\beta_{OWN}$ . The weaker distant-style comovements may be explained by the style-switching behaviour of noise traders. If style investors finance increases in own-style allocations (e.g. Large-cap Value) by selling securities in every other style (everything except large-cap Value), then it could explain the weaker distant-style comovements in misvaluation. Nevertheless, the misvaluation factors are successful in capturing the relevant style-based comovements in misvaluation as changes in misvaluation are not positively correlated with the distant-style factor, which could happen if I had omitted an important style category from the own-style.

Premiums predict changes in misvaluation negatively consistent with arbitrage activity and mean-reversion ( $\lambda_i < 0$  in Eq. (6)). As a benchmark, the  $\lambda_i$  coefficient should theoretically equal -1 if shocks to premiums fully revert over one period (i.e. when premiums are stationary). There is, however, a considerable degree of heterogeneity in the coefficient estimates with the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentiles at -0.85, 0.49, -0.26 respectively at the daily horizon.

Finally, it is also important to remember that the economic magnitudes documented here are conservative because with a twin-approach we can only identify the relative magnitude of excess comovements (the relative sensitivity to demand shocks  $\gamma_i^{ETF} - \gamma_i^{NAV}$  in Eq. (2)). The total amounts ( $\gamma_i^{ETF}$  and  $\gamma_i^{NAV}$ ) can be bigger if both ETFs and NAVs are affected by non-fundamental demands, which is likely to occur because ETFs may not have sufficient liquidity to absorb all of the liquidity-demand by short-term investors (to be discussed in section 6).

### 3.5.2 Results: sub-samples

Table 3-4 provides additional sub-sample results for commonality in misvaluation. If stale pricing is affecting the results, then we should expect the results to be mainly driven by small-cap ETFs whose underlying stocks may not always be actively traded. In contrast, I find that the degree of commonality in misvaluation is relatively similar across small-, mid- and large-cap

ETFs. The raw coefficient estimates for the own-style beta are 0.843, 0.828 and 0.957 for large-, mid and small-caps respectively. Similar patterns are also observed for the proportional impact of common misvaluation relative to raw returns ( $\beta_{OWN} * STD(R_{OWN,t}^{E-N}) / STD(R_{i,t}^{ETF})$ ): 4.52, 3.84 and 5.50 percent for large-, mid- and small-cap ETFs. Given the strength of the results for large- and mid-caps, stale pricing is unlikely to be a concern here. The somewhat stronger results for small-cap ETFs is also consistent with my conjecture that liquidity facilitates excess comovements because small-cap ETFs are on average the most liquid (relative to their underlying securities) and are therefore more likely to attract short-term investors (see Chapter 2).

[Table 3-4]

Among the various styles, the results are strongest for value ETFs both in regards to the own-style comovements ( $\beta_{i,OWN} = 1.037$ ), as well as the distant styles comovements ( $\beta_{i,DIST} = -0.173$ ). The results for growth ETFs are weaker than for value ETFs, but similar to blend funds. The weaker results for blend funds, despite their higher overall liquidity, is consistent with the idea that growth and value ETFs are more attractive to style investors because they represent pure plays on investment styles. The results for sector ETFs also hold strongly, although the comovements are slightly weaker than for non-sector funds ( $\beta_{i,OWN} = 0.774$ ) most likely due to the difficulty of constructing precise own-style misvaluation factor based on three distinct styles.

Finally, sub-sample results by time-period show that the own-style betas have generally increased over time from 0.590 in the period prior to the main sample (06/2002-12/2006), to 0.724 during the pre-financial crisis period (12/2006- 07/2008), to 0.884 during the crisis period (08/2008-03/2009) and to 0.799 in the post-crisis period (03/2009-12/2012). This pattern is consistent with a greater use of ETFs by short-term investors with non-fundamental demand as the overall liquidity of ETFs has increased over time (see Figure 2-1, in Chapter 2).

### 3.5.3 Robustness: Exposure to systematic risk

Differences in systematic risk between ETFs and their underlying portfolios might be able to explain the style-based commonality in misvaluation documented earlier. To investigate this

possibility, I regress ETF-NAV returns on the Fama and French 3-factors (*MKT*, *SMB* and *HML*<sup>35</sup>) and the lagged premium to control for mean-reversion in ETF-NAV returns.

[Table 3-5]

Table 3-5 shows that daily ETF-NAV returns are negatively and significantly exposed to the market factor (holds for the average ETF in every style); large-cap ETFs have a positive and significant exposure to *SMB*, small-caps have a negative and marginally significant exposure to *SMB*, while ETFs in every style are negatively and significantly exposed to *HML* (except for Growth ETFs). At the daily level the  $R^2$  increases by roughly 8 percent compared with the baseline model that only includes that lagged premium. A decomposition of  $R^2$  reveals that most of this increase is due to the negative exposure of ETF-NAV returns on the market factor. These results, particularly the uniformly negative exposure on the market factor, are not easy to reconcile with the style-based commonality in misvaluation documented earlier, nor do the results line up with the explanation that ETF returns are fundamentally more risky relative to NAV returns. Moreover, these findings disappear at lower return horizons.

Another possibility is that ETFs are differentially exposed to systematic liquidity risk, especially because there are large differences in liquidity between the ETF and its underlying portfolio. This story is, however, unlikely because recent evidence on the pricing of liquidity risk in U.S. stocks suggests that the characteristic liquidity premium has declined considerably over time and is priced only among the smallest stocks, while systematic liquidity is priced primarily among NASDAQ stocks (Ben-Rephael, Kadan and Wohl, 2013). In contrast, my results are not driven by small-cap ETFs. To formally investigate this issue I augment the Fama-French 3-factor model with the market-wide funding liquidity factor based on Hu, Pan and Wang (2013), which is available at the daily level from the author's website. In specification (b) I show that HPW's funding liquidity variable enters with a positive and significant coefficient for large-cap ETFs, but not for small-caps as we might have expected. As before, the results are insignificant at lower horizons. Thus, differences in systematic risk are unlikely to be able to explain the comovement patterns documented earlier among ETF-NAV returns.

---

<sup>35</sup> It is possible that ETF-NAV returns are correlated with *SMB* and *HML* if these return premiums are related to correlated non-fundamental demand. However, identification of this relationship is likely to be weak given that *SMB* and *HML* are not filtered from fundamental sources of risk.

### 3.6 SOURCE OF MISVALUATION: ETF OR NAV?

In this section I provide more direct evidence that ETF premiums are driven by misvaluation in the ETF leg (non-fundamental demand shocks) as opposed to in the NAV leg (slow diffusion of information). To illustrate the testable implications, suppose that the ETF  $i$  is hit by a positive *non-fundamental* demand shock that pushes its price above the underlying portfolio NAV value ( $PREM_{i,t} > 0$ ). If  $PREM_{i,t}$  truly reflects misvaluation of the ETF, then we should observe a reversal in the future returns of ETF  $i$  ( $R_{i,t+1}^{ETF} < 0$ ) with no impact on NAV returns ( $R_{i,t+1}^{NAV} = 0$ ).

The alternative hypothesis is that the initial demand shock was driven by positive *fundamental* news. In this case ETF  $i$  is correctly valued because the fundamental information is incorporated first into ETF prices, while its underlying portfolio NAV is incorrectly priced because it reacts more slowly, either because the high liquidity of ETFs attract fundamental traders, or due to stale pricing. The difference from before is that the price is correct for ETF  $i$  and its future returns remain unaffected ( $R_{i,t+1}^{ETF} = 0$ ). In contrast, future NAV returns will be positive as the NAV catches up to reflect the fundamental news already incorporated in the price of ETF  $i$  ( $R_{i,t+1}^{NAV} > 0$ ). To test these implications, I estimate the following regressions while controlling for lagged returns:

$$R_{i,t+1}^{ETF} = a_i^{ETF} + \sum_{k=0}^3 b_k^{ETF} PREM_{i,t-k} + \sum_{k=0}^3 c_k^{ETF} R_{t-k}^{ETF} + e_{i,t+1}^{ETF} \quad (8)$$

$$R_{i,t+1}^{NAV} = a_i^{NAV} + \sum_{k=0}^3 b_k^{NAV} PREM_{i,t-k} + \sum_{k=0}^3 c_k^{NAV} R_{t-k}^{NAV} + e_{i,t+1}^{NAV} \quad (9)$$

where  $R_{i,t+1}^{ETF}$  = ETF return measured on day  $t+1$  (over two consecutive trading days)  
 $R_{i,t+1}^{NAV}$  = NAV return measured on day  $t+1$   
 $PREM_{i,t}$  = ETF premium relative to NAV on day  $t$

If  $PREM_{i,t}$  reflects misvaluation for ETF  $i$ , then a future reversal in misvaluation implies that  $b_k^{ETF} < 0$ , while NAV returns remain unaffected  $b_k^{NAV} = 0$ . In contrast, the information diffusion story predicts that  $PREM_{i,t}$  reflects fundamental news already incorporated into the price of ETF  $i$  ( $b_k^{ETF} = 0$ ), but future returns of NAV will catch up with a lag ( $b_k^{NAV} > 0$ ). I also report the significance of the overall effect over a four day period ( $\sum b_k = 0$ ) using an F-test.

The main challenge in estimating regressions (8-9) is the dependence in residuals across funds because we have not accounted for common fundamental risk. To address this issue, I estimate (8-9) using pooled OLS with Driscoll and Kraay (1998) standard errors that are robust

to general forms of cross-sectional and time-series dependence. In unreported tests, I also verify that similar results hold if we aggregate (8-9) to the style level by taking an equally-weighted average of the LHS and RHS across all style categories (9 size-valuation styles for non-sector ETFs, 11 sector styles).

[Table 3-6]

Consistent with the reversal of non-fundamental demand shocks, the results in Table 3-6 show that premiums predict ETF returns negatively over a four day period with the overall effect being significant at the 5 percent level. Individually, the coefficients are negative for the first three lags, but significant only for the first two. The economic magnitude of the effect can be illustrated by using a one Std. Dev. shock to premiums: the overall effect ( $\sum b_k * STD(PREM_{i,t})$ ) equals a decline in ETF returns of 19.5 bps over a four day period. The magnitude of this effect is considerable when compared to the overall variability in changes in ETF misvaluation, which is 12 bps per day. In contrast, the overall effect of premiums on NAV returns over a four day period is insignificantly different from zero.

These results might also be affected by arbitrage activity, namely, from the price impact of buying the underpriced NAV and selling the overpriced ETF. In order to provide a more conservative test I investigate the net effect of premiums on future ETF returns relative to NAV returns ( $b_k^{ETF} - b_k^{NAV}$ ). If premiums reflect non-fundamental demand shocks in ETF prices that subsequently revert, then  $b_k^{ETF} - b_k^{NAV} < 0$ , whereas if premiums reflect fundamental demand shocks in ETF prices and which are subsequently incorporated into NAV prices, then  $b_k^{ETF} - b_k^{NAV} > 0$ . Surprisingly, the overall net effect ( $\sum(b_k^{ETF} - b_k^{NAV}) * STD(PREM_{i,t})$ ) is even more negative than before, at -30.1 bps, because premiums predict NAV returns with a *negative* sign on days two and three.

This finding is in contradiction with the information diffusion story predicts, but it can be explained in the context of non-fundamental demand shocks. Suppose investors trade sequentially, possibly because information (whether fundamental or non-fundamental) arrives sequentially. In this case, a positive non-fundamental demand shocks goes first to the most liquid ETFs. Once liquidity dries up in the most liquid ETF, it goes to the next most liquid and so on, until no more ETFs are liquid enough relative to their underlying basket, in which case the

demand goes to the underlying securities<sup>36</sup>. Thus, a positive premium reflects misvaluation in both the ETF and the NAV because both are hit by the demand shock, but the ETF is hit harder because it is more liquid. Consequently, both ETF and NAV prices are above their true fundamental values, in which case the returns of both must revert.

I also estimate regression (8-9) on sub-samples based on non-sector vs. sector, and for large-, mid- and small-cap funds. The results hold strongly for both non-sector and sector funds, although among the non-sector funds, only small-cap ETFs show significant evidence of return reversals following a positive premium. This finding is not only consistent with non-fundamental demand, but it also agrees with the conjecture that more liquid securities attract short-term investors with non-fundamental demand because small-cap ETFs have on average higher relative liquidity compared to either mid- or large-cap ETFs (Chapter 2).

### 3.7 CORRELATED DEMAND AND EXCESS COMOVEMENT IN RETURNS

The non-fundamentals-based view of excess comovement predicts that if the degree of commonality in demand shocks varies across securities or over time, then this variation should be positively related to the degree of commonality in misvaluation and the liquidity of the fund (Hypothesis 3). In order to test this hypothesis, we need empirical proxies for correlated demand shocks. In section 7.1, I propose two such measures. In section 7.2 I show that these demand shock proxies exhibit similar style-based comovements as do changes in ETF misvaluation. In section 7.3, I provide formal tests for Hypothesis 3.

#### 3.7.1 Measuring correlated demand shocks

To arrive at a proxy for abnormal demand shocks<sup>37</sup>, I build on the concept of abnormal trading activity. In the context of portfolio theory, turnover is a natural proxy for trading activity (Lo and Wang, 2000). Hence, I use the turnover of an ETF *relative* to its underlying basket of securities:

$$REL(TO)_{i,t} = \ln \left( TO_{i,t}^{ETF} / \sum_{k=1}^K w_{i,k,t} TO_{k,t}^{UND} \right) \quad (10)$$

---

<sup>36</sup> An alternative explanation is that liquidity demand by retail investors goes first to the most liquid ETF, then to the next most liquid, and all the way to the least liquid ETF, because trading in the underlying securities is always too costly. For institutional investors with the capacity to invest in the underlying securities, the demand would go to ETFs only as long as ETF liquidity is above the liquidity of the underlying securities.

<sup>37</sup> Here I do not attempt to differentiate between the various sources (i.e. sentiment, risk aversion, or liquidity needs)

where:  $TO_{i,t}^{ETF} = VOL_{i,t}^{ETF} / SHR_{i,t}^{ETF}$ , or the share volume divided by the number of shares outstanding for ETF  $i$  on day  $t$ .

$TO_{k,t}^{UND}$  = turnover of underlying security  $k$  on day  $t$ .

$w_{i,k,t}$  = dollar-weight invested by ETF  $i$  in security  $k$  at the end of day  $t$

Lo and Wang (2000) use a similar measure of portfolio turnover. Higher numbers for  $REL(TO)$  indicate that the ETF is more actively traded relative to its underlying basket, presumably because the ETF attracts high-turnover investors. To arrive at a measure for unexpected shocks to relative turnover, I use the residual from an AR(1) model, which I denote by  $\Delta REL(TO)$ . Unexpected increases in  $\Delta REL(TO)$  may reflect either an unexpected increase in ETF, or NAV trading activity. The shocks are, however, most likely to come from the ETF rather than the NAV because the correlation between  $\Delta REL(TO)$  and a similarly constructed measure for shocks to ETF turnover is roughly 0.9 (see also Chapter 2, Table 2-2).

To investigate commonality in relative turnover, I adopt the same approach that I used for ETF-NAV returns. Specifically, I regress shocks to relative trading activity on the equally-weighted shock to relative trading activity of other funds in ETF  $i$ 's own or distant styles (as defined in the previous section):

$$\Delta REL(TO)_{i,t} = \alpha_i + \sum_{j=-1}^{+1} \left( \beta_{i,j,O} \Delta REL(TO)_{OWN,t+j} + \beta_{i,j,DI} \Delta REL(TO)_{DIST,t+j} \right) + u_{i,t} \quad (11)$$

The one-day leading and lagged terms are meant capture any lagged adjustment in commonality (Chordia, Roll and Subrahmanyam, 2000). Correlated demand at the style level implies positive concurrent own-style betas ( $\beta_{i,O}$ ). One caveat is that I cannot rule out comovements across styles ( $\beta_{i,DI} > 0$ ) because  $\Delta REL(TO)$  may also capture fundamental demand shocks. Nevertheless, I would expect to find stronger own- than distant-style comovements if the non-fundamental style component is strong.

As another measure of correlated demand, I use the degree of commonality in relative liquidity. There is an extensive literature documenting that liquidity comoves across stocks. The demand-side view argues that commonality in liquidity arises because of correlated trading activity (Chordia, Roll and Subrahmanyam, 2000; Karolyi, Lee and Van Dijk, 2012), demand by institutional owners (Kamara, Lou and Sadka, 2008), by investor sentiment (Huberman and Halka, 2001) or by the price impact of correlated liquidity needs (Greenwood and Thesmar,

2011). In this case we can view commonality in ETF liquidity as a proxy for correlated demand. The supply-side view provides a different interpretation. In this case liquidity commonality is explained by the funding constraints of financial intermediaries. Several theoretical models predict that commonality in liquidity, via illiquidity spirals or feedback loops, increases during periods when arbitrage capital is limited<sup>38</sup>. However, as we shall see in section 7.2, the results are more consistent with the demand-side view of liquidity commonality.

To measure relative liquidity, I use the difference between the (log of) Amihud's price impact<sup>39</sup> for the underlying portfolio and the ETF:

$$REL(PI)_{i,t} = \log \left( \sum_{k=1}^K w_{i,k,t} \frac{|R_{k,t}^{UND}|}{DVOL_{k,t}^{UND}} / \frac{|R_{i,t}^{ETF}|}{DVOL_{i,t}^{ETF}} \right) \quad (12)$$

where  $R_{i,t}^{UND}$  = mid-quote return (in %) for security  $k$  held by ETF  $i$ , on trading day  $t$   
 $DVOL_{i,t}^{UND}$  = dollar volume (in \$millions) for security  $k$ , on trading day  $t$

Amihud's Price Impact ( $PI$ ) has been widely used in the literature. Hasbrouck (2009) reports that, "among the daily proxies, the Amihud measure is most strongly correlated with the TAQ-based price impact coefficient" (p. 1459). Amihud's measure is also endorsed by several other papers as good proxy for price impact; others have used it to study commonality in liquidity<sup>40</sup>. A similar measure of portfolio liquidity has been used by Idzorek, Xiong and Ibbotson (2012). Having defined  $REL(PI)$ , parallel calculations are done to compute measures of commonality with  $REL(TO)$  replaced by  $REL(PI)$  in Eq. (10). The data for portfolio weights comes from Morningstar Direct<sup>41</sup>. For a more detailed description and summary statistics of these variables, see Chapter 2.

According to the non-fundamentals based view of excess comovement there is a positive relationship between the degree of commonality in misvaluation and turnover/liquidity because both are driven by a common factor, namely correlated demand shocks. It is important to emphasize that this prediction does *not* imply that ETF-NAV returns can be explained by relative

<sup>38</sup> see Karolyi, Lee and Van Dijk (2012) for an extensive list of references.

<sup>39</sup> Daily observations of the price impact ratio above the 99.5<sup>th</sup> percentile of the sample have been discarded as in Amihud (2002). Similar results obtain if I use the CRSP-based quoted spreads to measure liquidity.

<sup>40</sup> Lesmond (2005), Goyenko, Holden and Trzcinka (2009), Fong, Holden, and Trzcinka (2010) endorse Amihud, while Karolyi, Lee and Van Dijk (2012) and Kamara, Lou and Sadka (2008) use Amihud for liquidity commonality.

<sup>41</sup> Since my holdings data for the underlying holdings of an ETF is generally at the monthly level, the implicit assumption is that changes in weights only reflect changes in market values of the constituents.

turnover/liquidity (or by the own-style turnover/liquidity factors). In particular, commonality in turnover/liquidity tends to be high during market downturns when volatility is high and when returns are extremely low (Karolyi, Lee and Van Dijk, 2012), while liquidity is high in the opposite state and turnover can be high in either extreme state.

The theoretical model by Cherkes, Sagi and Stanton (2008) does, however, predict a positive relationship between the return difference of twin securities and their liquidity difference because premiums reflect a trade-off between liquidity and expense ratios. Their prediction can explain the style-based commonality in misvaluation documented previously only if changes in relative liquidity are correlated across ETFs at the style level, and if such common changes in liquidity can explain ETF-NAV return differences. In unreported tests, I investigate this issue by regressing changes in premiums (i.e. ETF-NAV returns) on shocks to relative turnover and relative liquidity (Eq. (10) and (12)) and the own- and distant-style factors for turnover/liquidity. The liquidity measures are consistently insignificant.

### 3.7.2 Results: commonality in demand shocks

Table 3-7 presents the results for correlated demand shocks, as estimated from Eq. (11) for shocks to relative turnover (Panel A), or relative liquidity (Panel B). I report the following results: average and median values for the concurrent, lagged, lead, sum coefficients and  $R^2$ ; the percentage of funds with positive coefficients, the percentage of funds with positive and significant coefficients, negative and significant coefficients. Test of statistical significance for the average (median) coefficient is based on the cross-sectional t-statistic (sign-test) similar to Chordia, Roll and Subrahmanyam (2000) and Brockman *et al.* (2009). The results show that shocks to relative turnover  $\Delta REL(TO)$  comove positively and significantly across ETFs in the same style both at the mean and the median. More than 93 percent of the concurrent own-style betas ( $\beta_{i,O}$ ) are positive and (individually) significant at least at the 5 % level. Although shocks to relative trading activity also comove across distant styles ( $\beta_{i,DI} > 0$ ), the magnitude of the distant-style betas are less than a third as large as the own-style betas.

[Table 3-7]

The results for relative liquidity are even stronger: shocks to relative liquidity exhibit positive and significant own-style comovements (in 99 percent of cases), while the distant-style comovements are significantly *negative* ( $\beta_{i,DI} < 0$ ). This is consistent with the earlier results for

commonality in misvaluation and with the prediction by Barberis and Shleifer (2003) that increases in own-style allocations are at least partly financed by decreases in distant-style allocations. Moreover, these style-based comovements in turnover and liquidity are more consistent with demand than supply-side explanations given that the theoretical effects behind the latter (illiquidity spirals and feedback loops) are generally described as a market-wide phenomenon.

Overall, the results in this section highlight that proxies for demand shocks (relative turnover and price impact) exhibit similar style-based comovements as do ETF-NAV returns. Similar results are also obtained when turnover/liquidity shocks are measured at the ETF level (ETF turnover or price impact instead of relative turnover or relative price impact), suggesting that commonality in demand shocks is mainly coming from the ETF, rather than the NAV leg.

### **3.7.3 Explaining the amount of commonality in misvaluation**

Hypothesis 3 predicts a link between the degree of commonality in misvaluation, commonality in demand shocks and the level of fund liquidity. In this context, what is the appropriate measure of the *degree* of commonality? The existing literature mainly uses the regression  $R^2$  (e.g. Morck, Yeung and Yu, 2000; Hameed, Kang and Viswanathan, 2010; Karolyi, Lee and Van Dijk, 2012), although the beta coefficient is also used (Kamara, Lou and Sadka, 2008).

I use the  $R^2$  measure for three reasons. First, the beta is sensitive to scaling effects that arise from differences in factors and their volatilities (i.e. the beta denominator) across ETF styles. Second, it is difficult to make cross-sectional comparisons of betas in short samples due to large cross-sectional differences in the Std. Dev. of ETF-NAV returns, which is also directly related to arbitrage costs (funds with higher arbitrage costs have more volatile ETF-NAV returns). The  $R^2$ -measure does not suffer from these problems as it is a function of both the variance of the dependent variable and the factors. Another interpretation of the  $R^2$ -measure is that it captures the proportion of common vs. idiosyncratic risk, in which case it is less sensitive to total misvaluation. This metric is also suitable for testing the hypothesis that commonality in misvaluation is stronger for more liquid ETFs. In unreported robustness tests I verify that the main results continue to hold for beta coefficients.

Following Karolyi, Lee and Van Dijk (2012), I estimate the regression  $R^2$  (labelled  $R_{ret,m}^2$ ) from Eq. (6) every month  $m$  on daily data (at least 15 non-missing observations per month). Since regression (6) also controls for the lagged premium, I decompose the model  $R^2$  as:

$$R_m^2 = \frac{COV(R_{i,t}^{E-N}, \lambda_i P_{i,t-1}^{E-N})}{VAR(R_{i,t}^{E-N})} + \frac{COV(R_{i,t}^{E-N}, \beta_{i,OWN} R_{OWN,t}^{E-N})}{VAR(R_{i,t}^{E-N})} + \frac{COV(R_{i,t}^{E-N}, \beta_{i,DI} R_{DI,t}^{E-N})}{VAR(R_{i,t}^{E-N})} \quad (13)$$

and use the sum of the last two normalized covariance terms, denoted  $R_{ret,m}^2$ , to measure the degree of commonality in misvaluation.  $R_{ret,m}^2$  can be interpreted as the fraction of the model  $R^2$  attributable to the own- and distant-style factors (Graham, Li and Qiu, 2013). Similarly, I measure commonality in relative turnover/liquidity shocks from the fraction of model  $R^2$  attributable to the concurrent own and distant-style factors from Eq. (11). The degree of commonality in relative turnover and price impact is denoted by  $R_{rTO,m}^2$  and  $R_{rPI,m}^2$  respectively.

To investigate the relationship between commonality in misvaluation, commonality in demand shocks and the attractiveness of an ETFs liquidity characteristics (both direct and indirect), I estimate the following regression using pooled OLS:

$$R_{ret,m}^2 = a + b_1 R_{dem,m-1}^2 + b_2 LIQ_{i,m-1} + b_3 Fund_{i,m-1} + b_4 Macro_m + FE + e_{i,m} \quad (14)$$

where  $R_{dem,m-1}^2$  = commonality in demand shocks based on  $R_{rTO,m}^2$  or  $R_{rPI,m}^2$

$LIQ_{i,t-1}$  = ETF & underlying portfolio liquidity during month  $t-1$

$Fund_{i,m-1}$  = vector of fund characteristics (incl. indirect liquidity characteristics)

$Macro_m$  = vector of macro variables

$FE$  = fixed effects: year, month, sector, style and/or fund

All independent variables are lagged by one month to mitigate concerns about reverse causality.

As a direct and salient measure of liquidity, I use the monthly average quoted spread for ETF  $i$ :

$$QSPR_{i,m}^{ETF} = -1 * \ln \left( \frac{1}{N_m} \sum_{t=1}^{N_m} 100 * \frac{ASK_{i,t}^{ETF} - BID_{i,t}^{ETF}}{(ASK_{i,t}^{ETF} + BID_{i,t}^{ETF}) / 2} \right) \quad (15)$$

where:  $ASK_{i,t}^{ETF}$  and  $BID_{i,t}^{ETF}$  = CRSP ask and bid price at the close on trading day  $t$  for ETF  $i$

$N_m$  = nr. of trading days in calendar month  $m$

I use the log-transformation to mitigate the impact of outliers and to deal with the apparent non-stationarity in the data. I estimate the portfolio quoted spread by dollar-weighting the monthly

quoted spread of each security included in the ETF's basket ( $QSPR_{i,m}^{UND} = w_{i,k,m} QSPR_{k,m}^{UND}$ ). According to Chung and Zhang (2014), the CRSP-based spread is highly correlated with the (more accurate) TAQ spread in the cross-section, which is the dimension of primary interest. I also include the expense ratio because it is a salient cost for retail investors (Grinblatt *et al.*, 2014). Another implicit measure of liquidity is total misvaluation, measured by the absolute value of the monthly average premium (Eq. (4)), because funds with high misvaluation have high arbitrage costs.

Arbitrage activity should be negatively related to the degree of commonality in misvaluation because it is associated with greater pricing efficiency. To measure arbitrage activity, I use share creation/redemption activity, defined as in Chapter 2:

$$CREATE_{i,m} = \log \left[ 1 + \frac{1}{N_m} \sum_{d=1}^{N_m} \frac{|SHR_{i,d} - SHR_{i,d-1}|}{SHR_{i,d-1}} \right] \quad (16)$$

where:  $SHR_{i,d}$  = shares outstanding for ETF  $i$  on day  $d$ ;  
 $N_t$  = number of trading days in calendar month  $t$ .

Commonality in misvaluation should be greater when market-wide arbitrage costs are high because they leave more “room” for excess comovement (Kumar and Lee, 2006; Kumar and Spalt, 2013). I use the funding liquidity factor by Hu, Pan and Wang (2013), which is based on price deviations between on-the-run and off-the-run Treasury securities, averaged across a wide range of maturities. Market volatility is also an important determinant of the risk to market makers of maintaining inventories of their securities (Chordia, Roll, and Subrahmanyam, 2000), and changes in market volatility can cause changes in inventories and create correlated institutional trading. Market volatility is also related aggregate uncertainty in financial markets either via higher transaction costs or lower funding liquidity (i.e., less capital is devoted to ETF arbitrage) as in Brunnermeier and Pedersen (2009). In either case, the prediction is that commonality in misvaluation should be positively related to market volatility, which I proxy for by the Std. Dev. of NAV returns.

[Figure 3-1]

Before I discuss the regression results, I illustrate the time-series dynamics for the cross-sectional mean degree of commonality in misvaluation ( $R_{ret,m}^2$ ) in Figure 3-1. Panel A depicts

$R_{ret,m}^2$  across three terciles based on the degree of commonality in relative liquidity (via price impact) in the prior month ( $R_{rPI,m-1}^2$ ), while in Panel B, I instead use three terciles based on relative quoted spreads ( $QSPR_{i,m-1}^{ETF} - QSPR_{i,m-1}^{UND}$ ). We can see that ETFs in the top tercile of commonality in relative price impact have a higher degree of commonality in misvaluation in 99 % of quarters (93 % are significant at the 5 % level) relative to ETFs in the lowest tercile. Moreover, ETFs in the highest tercile of relative quoted spreads have stronger commonality in misvaluation in 98 % of the quarters (70 % are significant at the 5 % level). These preliminary findings agree with the conjecture that commonality in misvaluation is higher when commonality in demand shocks is high, and for ETFs with more desirable liquidity characteristics.

[Table 3-8]

The results in Table 3-8 show that commonality in misvaluation is positively and significantly related to commonality in demand shocks (via  $R_{rTO}^2$  and  $R_{rPI}^2$ ). When both measures for commonality in demand shocks are included simultaneously,  $R_{rPI}^2$  remains highly significant while  $R_{rTO}^2$  is only marginally significant. This is not surprising because the two measures are designed to capture the same thing. Moreover, commonality in misvaluation is significantly higher for ETFs with more attractive liquidity characteristics (lower quoted spreads, expense ratios and total misvaluation), for less liquid underlying portfolios, and during times when arbitrage is limited (high market volatility, low funding liquidity). Similar results are obtained in sub-samples of non-sector vs. sector funds, or if we control for time or ETF fixed effects in addition to style fixed effects.

Overall, the findings in this section show that the degree of commonality in misvaluation is stronger for ETFs with more correlated demand shocks and more desirable liquidity characteristics, which supports my conjecture that excess comovements in ETF returns are driven by correlated non-fundamental demand and facilitated by investors with short horizons that are attracted to ETFs because of their high liquidity.

### 3.8 SUMMARY AND CONCLUSIONS

This study analyzes whether the returns of Exchange Traded Funds (ETFs) comove excessively with other ETFs in similar investment styles. My conjecture is that, due to the ease of investing in ETFs and because of their high liquidity, ETFs attract a clientele of short-term investors who

are more exposed to common non-fundamental demand shocks at the style level relative to the investor in the ETFs underlying baskets.

In order to identify excess comovements, I look for common factors in changes in misvaluation (ETF-NAV returns); in order to attribute this misvaluation to the ETF leg, I study return reversals following shocks to misvaluation; and in order to link the degree of excess comovement to common non-fundamental demand shocks and to liquidity clientele differences, I investigate how the degree of commonality in misvaluation is related to the degree of commonality in turnover/liquidity (proxy for demand shocks) and ETF characteristics that are attractive to investors with short-term trading needs.

My findings indicate that there is significant style-based commonality in misvaluation: changes in misvaluation comove positively across ETFs in similar styles, and negatively with ETFs in distant styles. Although the importance of common misvaluation is found to persist across daily, weekly and monthly horizons, the economic magnitude of these shocks declines with the return horizon consistent with arbitrage. The source of misvaluation is found to be the ETF since current ETF premiums negatively predict future ETF returns consistent with temporary demand shocks in the ETF prices. Such reversals are strongest for small-cap ETFs, which is consistent with the conjecture that more liquid securities (small-cap ETFs have the highest relative liquidity) attract short-term investors with non-fundamental demand.

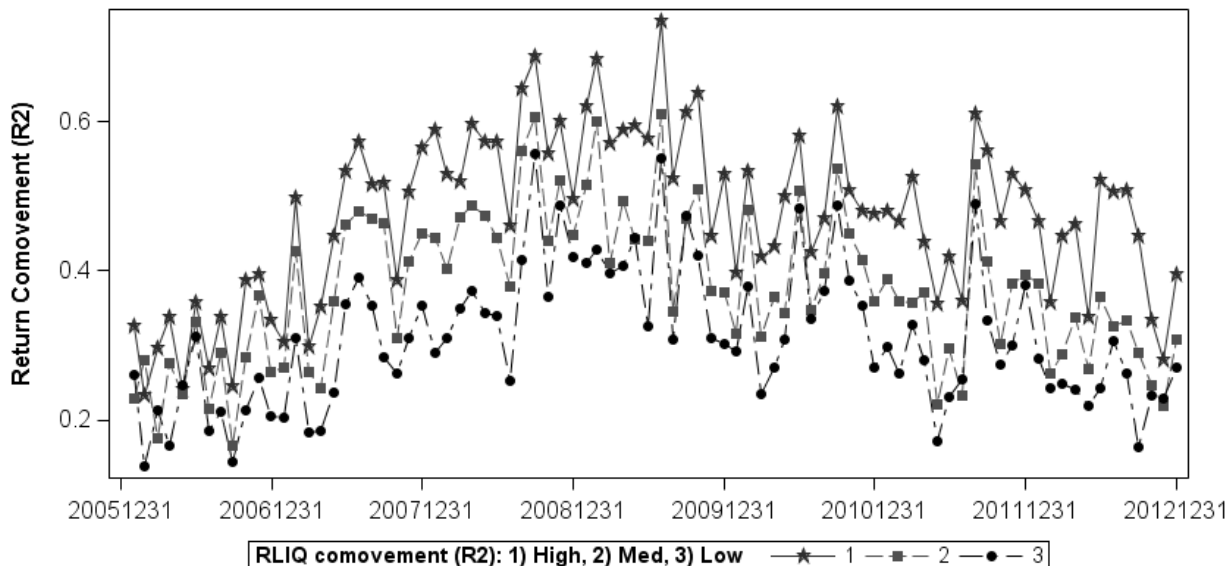
Moreover, shocks to the abnormal trading activity of ETFs and to the liquidity of ETFs (both relative to their underlying baskets) exhibit similar style-based commonality. These proxies for correlated demand shocks also predict the future amount of commonality in ETF misvaluation. In accordance with liquidity being a factor in inducing clientele differences between ETFs and their underlying portfolio, I show that the degree of commonality in misvaluation is higher among ETFs more desirable liquidity characteristics.

My overall conclusion is that the excess comovement in ETF returns is mainly driven by the correlated non-fundamental demand of a liquidity-based clientele. Thus, more liquid securities may encourage trading for reasons unrelated to fundamentals, which can at times be detrimental for pricing efficiency.

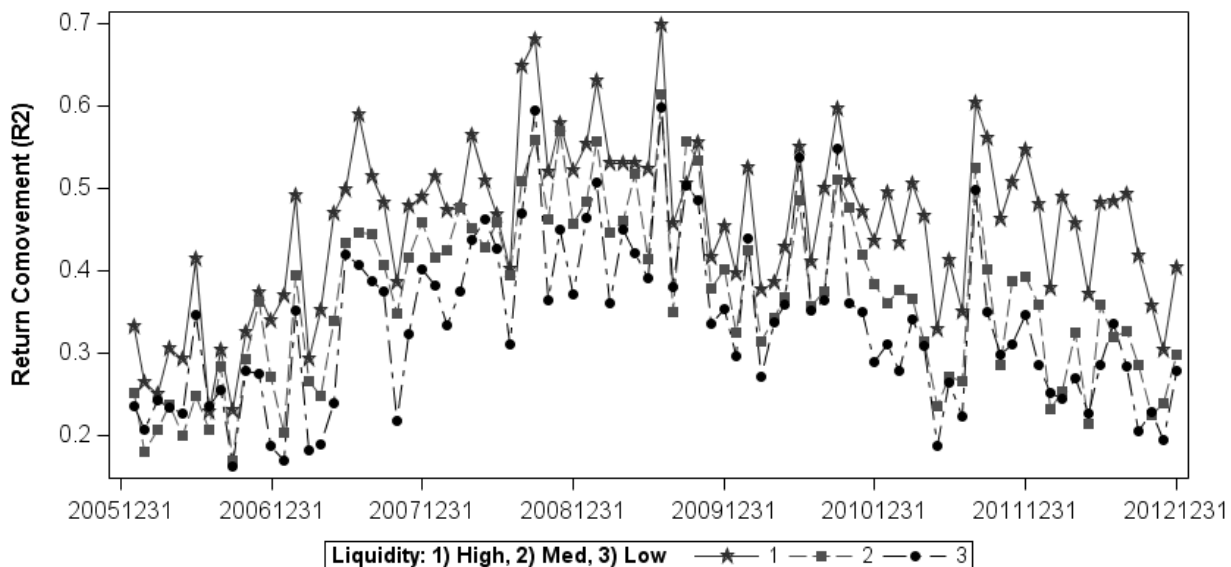
**Figure 3- 1: Commonality in misvaluation, commonality in demand shocks and the level of liquidity**

The cross-sectional mean degree of commonality in misvaluation is plotted over time for three groups. Terciles are based on commonality in relative price impact ( $R_{rlq,q-1}^2$ ) and the level of relative quoted spreads in Panels A and B respectively. The degree of commonality is based on the model  $R^2$ , or the fraction of the variation in the dependent variable (ETF-NAV returns, relative turnover or liquidity) attributable to the own and distant style factors.

**Panel A: Cross-sectional mean of  $R_{ret,q}^2$  by  $R_{rlq,q-1}^2$  tercile**



**Panel B: Cross-sectional mean of  $R_{ret,q}^2$  by  $(QSPR_{i,m-1}^{ETF} - QSPR_{i,m-1}^{UND})$  tercile**



**Table 3- 1: Snapshot of ETF statistics**

This table reports the number of ETFs and the total Assets under Management (AUM) at the beginning, the middle and at the end of the sample. The statistics are reported by Morningstar 3-by-3 style classification (size and valuation) for ETFs with a core style, and by Morningstar industry styles for sector ETFs.

Category	Nr. of ETFs			AUM (in \$millions)		
	01/2006	06/2009	12/2012	01/2006	06/2009	12/2012
<b>Core styles</b>	<b>53</b>	<b>73</b>	<b>79</b>	<b>168,391</b>	<b>210,116</b>	<b>425,491</b>
Large-Value	8	12	11	16,042	20,398	43,041
Large-Blend	12	16	21	81,368	105,540	220,168
Large-Growth	8	11	12	31,840	33,355	64,670
Mid-Value	3	5	5	4,589	4,020	7,373
Mid-Blend	4	6	6	6,906	16,174	34,492
Mid-Growth	4	5	5	4,008	4,404	7,854
Small-Value	5	4	3	5,228	4,178	6,661
Small-Blend	4	8	10	13,874	16,321	31,925
Small-Growth	5	6	6	4,537	5,727	9,307
<b>Sector</b>	<b>42</b>	<b>83</b>	<b>85</b>	<b>31,438</b>	<b>47,428</b>	<b>114,520</b>
Communication	2	2	2	551	650	936
Cons. Cyclical	3	7	7	759	3,158	9,348
Cons. Defens.	3	4	4	1,334	2,756	7,258
Energy	4	9	9	6,657	8,067	12,877
Financial	4	11	13	2,383	8,756	14,098
Real Estate	4	10	10	5,012	5,652	26,087
Health Care	4	11	11	5,216	5,448	11,498
Industrials	4	8	8	1,314	2,188	5,953
Materials	3	5	5	1,241	2,406	4,675
Technology	8	12	12	3,851	5,562	14,239
Utilities	3	4	4	3,120	2,788	7,552
All	95	156	164	199,829	257,545	540,011

**Table 3- 2: Descriptive Statistics**

$PREM_{i,t} (R_i^{E-N})$  is the level of (change in) premium calculated as the log-price (return) differential between ETF  $i$  and its underlying portfolio Net Asset Value ( $NAV$ ). Both levels and changes in premiums are reported in percentage. Closing and mid-point refers to premiums calculated using closing or mid-point prices/returns. Summary statistics are calculated using daily data, over the 01/2006 to 12/2012 period. Statistics are reported by Morningstar's 3-by-3 style classification for funds with a core (non-sector) style, and by Morningstar industry style for sector funds.

Variable	Style	Mean	Median	Std. Dev.	1 %	99 %
Closing: $PREM_{i,t} = \ln(P_i^{ETF} / P_i^{NAV})$	All funds	-0.005	0.000	0.163	-0.553	0.522
	All funds	-0.005	-0.003	0.091	-0.268	0.271
	<b>By size</b>					
	Large	-0.015	-0.013	0.115	-0.372	0.307
	Mid	-0.007	-0.007	0.083	-0.220	0.248
	Small	0.001	0.000	0.080	-0.198	0.252
	<b>By valuation</b>					
	Blend	-0.003	0.000	0.088	-0.257	0.262
	Value	-0.004	0.000	0.092	-0.277	0.289
	Growth	-0.007	-0.007	0.093	-0.277	0.270
<b>By core vs. sector</b>						
Core	-0.005	-0.003	0.091	-0.268	0.271	
Sector	-0.005	-0.007	0.121	-0.302	0.361	
Closing: $R_i^{E-N} = R_i^{ETF} - R_i^{NAV}$	All funds	0.000	0.000	0.224	-0.730	0.710
	All funds	0.000	0.000	0.120	-0.371	0.356
	<b>By size</b>					
	Large	0.000	0.000	0.150	-0.457	0.444
	Mid	0.000	0.000	0.119	-0.338	0.334
	Small	0.000	0.000	0.103	-0.311	0.301
	<b>By valuation</b>					
	Blend	0.000	0.000	0.113	-0.350	0.337
	Value	0.000	0.000	0.121	-0.388	0.373
	Growth	0.000	0.000	0.127	-0.383	0.371
<b>By core vs. sector</b>						
Core	0.000	0.000	0.120	-0.371	0.356	
Sector	0.000	0.000	0.169	-0.463	0.465	
Mid-point: $PREM_{i,t} = \ln(P_i^{ETF} / P_i^{NAV})$	All funds	-0.005	0.000	0.163	-0.553	0.522
	All funds	-0.005	-0.003	0.091	-0.268	0.271
	<b>By size</b>					
	Large	-0.015	-0.013	0.115	-0.372	0.307
	Mid	-0.007	-0.007	0.083	-0.220	0.248
	Small	0.001	0.000	0.080	-0.198	0.252
	<b>By valuation</b>					
	Blend	-0.003	0.000	0.088	-0.257	0.262
	Value	-0.004	0.000	0.092	-0.277	0.289
	Growth	-0.007	-0.007	0.093	-0.277	0.270
<b>By core vs. sector</b>						
Core	-0.005	-0.003	0.091	-0.268	0.271	
Sector	-0.005	-0.007	0.121	-0.302	0.361	
Mid-point: $R_i^{E-N} = R_i^{ETF} - R_i^{NAV}$	All funds	0.000	0.000	0.224	-0.730	0.710
	All funds	0.000	0.000	0.120	-0.371	0.356
	<b>By size</b>					
	Large	0.000	0.000	0.150	-0.457	0.444
	Mid	0.000	0.000	0.119	-0.338	0.334
	Small	0.000	0.000	0.103	-0.311	0.301
	<b>By valuation</b>					
	Blend	0.000	0.000	0.113	-0.350	0.337
	Value	0.000	0.000	0.121	-0.388	0.373
	Growth	0.000	0.000	0.127	-0.383	0.371
<b>By core vs. sector</b>						
Core	0.000	0.000	0.120	-0.371	0.356	
Sector	0.000	0.000	0.169	-0.463	0.465	

**Table 3- 3: Style-based commonality in misvaluation**

This table reports results from estimating the following regression, fund-by-fund:

$$R_{i,t}^{E-N} = \alpha_i + \lambda_i PREM_{i,t-1} + \beta_{i,OWN} R_{OWN,i,t}^{E-N} + \beta_{i,DIST} R_{DIST,i,t}^{E-N} + e_{i,t}$$

where  $R_i^{E-N}$  is the ETF-NAV return difference (or equivalently the change in premium,  $\Delta P_i^{E-N}$ ) and  $PREM_i$  is the ETF premium.  $R_{OWN}^{E-N}$  is the own-style misvaluation factor, which is based on the average ETF-NAV return of other ETFs in matching styles: equal weight is given to funds that match both style dimensions (size and valuation), and half the equal weight to funds that are in adjacent styles (if ETF  $i$  is Large-Value, then adjacent styles include Mid-Value and Large-Blend). For ETFs belonging to sector styles, equal weight is given to funds in the same industry, half the equal weight to funds in the same size and valuation category and one fourth of the equal weight to funds in adjacent core styles.  $R_{DIST}^{E-N}$  is the distant-style misvaluation factor and it is based on the average ETF-NAV return of other ETFs in opposite styles: equal weight is given to funds that are in opposite styles (both size and valuation) and half the equal weight to funds that are in styles adjacent to the opposite style. For instance, if ETF  $i$  is Large-Value, then Small-Growth is the opposite style (equal-weight) and Mid-Growth and Small-Blend (half the equal weight) are in styles adjacent to Small-Growth. For mid-cap funds we consider large and small to be equally distant. Blend funds (neither value, nor growth) are matched only by their size category. Hence, the distant style of Mid-Value is Small-Growth and Large-Growth. The t-statistic for the average coefficient is adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). Impact (basis points) is the impact of a 1 Std. Dev. increase in the RHS variable on the dependent variable, while impact  $[\%STD(R_{i,t}^{E-N})]$  and impact  $[\%STD(R_{i,t}^{ETF})]$  are the impact in basis points scaled by the Std. Dev. of the dependent variable and raw ETF return respectively. Coefficients and t-statistics are also reported at the 5th, 50th and 95th percentile of the cross-sectional distribution. In this case the t-statistics are based on heteroskedasticity-robust standard errors. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 %.

	Daily horizon			Weekly horizon			Monthly horizon		
	$\lambda_i$	$\beta_{i,OWN}$	$\beta_{i,DIST}$	$\lambda_i$	$\beta_{i,OWN}$	$\beta_{i,DIST}$	$\lambda_i$	$\beta_{i,OWN}$	$\beta_{i,DIST}$
Average coefficient	-0.523*** (26.44)	0.819*** (26.65)	-0.071*** (2.61)	-0.567*** (14.59)	0.783*** (13.34)	-0.057 (1.13)	-0.580*** (10.51)	0.737*** (8.46)	-0.025 (0.34)
<b>Economic significance:</b>									
Impact [basis points]:	-5.43	6.99	-0.63	-5.47	7.54	-0.55	-6.56	7.97	-0.26
Impact $[\%STD(R_{i,t}^{E-N})]$	-38.61	55.73	-3.72	-37.34	56.22	-4.38	-38.06	45.91	-0.39
Impact $[\%STD(R_{i,t}^{ETF})]$	-2.94	4.01	-0.33	-1.56	2.29	-0.19	-1.01	1.27	-0.02
<b>Distribution of coefficients:</b>									
5th percentile	-0.848*** (-14.48)	0.314* (1.91)	-0.399*** (-4.48)	-0.950*** (-8.73)	0.158 (0.68)	-0.531*** (-3.15)	-1.119*** (-7.27)	-0.210 (-0.36)	-0.853*** (-2.97)
50th percentile	-0.488*** (-8.77)	0.792*** (9.54)	-0.071 (-1.00)	-0.560*** (-4.50)	0.783*** (4.37)	-0.050 (-0.41)	-0.568*** (-3.01)	0.720*** (2.76)	-0.019 (-0.08)
95th percentile	-0.258*** (-3.96)	1.403*** (24.61)	0.234*** (2.92)	-0.195 (-1.48)	1.376*** (13.53)	0.462*** (2.22)	-0.008 (-0.03)	1.497*** (8.72)	0.601*** (2.41)
Average R <sup>2</sup>	0.64			0.67			0.62		
N obs	260,451			53,764			13,664		

**Table 3- 4: Sub-sample results for commonality in misvaluation**

This table reports results from estimating the following regression on daily data, fund-by-fund:

$$R_{i,t}^{E-N} = \alpha_i + \lambda_i PREM_{i,t-1} + \beta_{i,OWN} R_{OWN,i,t}^{E-N} + \beta_{i,DI} R_{DIST,i,t}^{E-N} + e_{i,t}$$

where  $R_i^{E-N}$  is the ETF-NAV return difference (or equivalently the change in premium,  $\Delta P_i^{E-N}$ ) and  $P_i^{E-N}$  is the ETF premium.  $R_{OWN}^{E-N}$  is the own-style misvaluation factor, which is based on the average ETF-NAV return of other ETFs in matching styles.  $R_{DIST}^{E-N}$  is the distant-style misvaluation factor and it is based on the average ETF-NAV return of other ETFs in opposite styles. For further details on the construction of these factors and the own vs. distant style classification, see Table 3. The beta coefficients reported are average betas across all funds in a given style. The t-statistic for the average coefficient is adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). Impact  $[\%STD(R_{i,t}^{E-N})] = 100 * \beta_{OWN,i} * STD(R_{OWN}^{E-N})/STD(R_{i,t}^{E-N})$ , or in words, the impact of a 1 Std. Dev. increase in the own-style factor as a percentage of the Std. Dev. of the dependent variable. Similarly, Impact  $[\%STD(R_{i,t}^{ETF})] = 100 * \beta_{OWN,i} * STD(R_{OWN}^{E-N})/STD(R_{i,t}^{ETF})$ , or the impact of a 1 Std. Dev. increase in the own-style factor as a percentage of the Std. Dev. of raw ETF returns. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 %.

Sub-sample	$\beta_{i,OWN}$	[t-stat]	Impact [ $\%STD(R_{i,t}^{E-N})$ ]	Impact [ $\%STD(R_{i,t}^{ETF})$ ]	$\beta_{i,DIST}$	[t-stat]	Avg. R <sup>2</sup>	N obs
All	0.819	[26.65]	55.73	4.01	-0.071	[2.61]	0.644	260,451
Diversified non-sector								
Large	0.843	[37.82]	65.17	4.52	-0.039	[3.51]	0.702	66,255
Mid	0.828	[19.96]	65.22	3.84	-0.083	[2.19]	0.680	28,613
Small	0.957	[49.05]	68.39	5.50	-0.088	[2.86]	0.728	31,197
Blend	0.800	[42.76]	62.81	4.38	0.000	[1.32]	0.717	52,821
Value	1.037	[35.14]	75.56	5.54	-0.173	[6.86]	0.725	34,475
Growth	0.805	[32.09]	61.41	4.09	-0.071	[3.59]	0.665	38,769
Sector	0.774	[18.99]	46.22	3.44	-0.081	[2.13]	0.588	134,881
Time-period: 01/2002 to 12/2006								
All	0.590	[8.60]	31.82	2.47	-0.028	[0.55]	0.553	65,003
Time-period: 12/2006 to 07/2008								
All	0.724	[21.68]	47.16	3.54	-0.042	[1.51]	0.724	82,927
Time-period: 08/2008 to 03/2009								
All	0.884	[15.62]	63.17	5.34	-0.102	[1.84]	0.673	25,300
Time-period: 03/2009 to 12/2012								
All	0.799	[30.27]	54.94	3.04	-0.070	[3.37]	0.660	152,719

**Table 3- 5: Differences in systematic risk?**

This table reports regressions of ETF-NAV returns on the lagged level of premium to control for mean-reversion, *MKT*, *SMB*, *HML* and funding liquidity (NOISE) factors, estimated fund-by-fund using all available observations. T-statistics for the mean are adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % level.  $\Delta R^2$  is the average improvement in R-squared compared to the model where ETF-NAV returns are only regressed against the lagged level of premium.

Factor	By style	Daily horizon		Weekly horizon		Monthly horizon	
		(a)	(b)	(a)	(b)	(a)	(b)
MKT	All	-0.0193***	-0.0197***	-0.0050**	-0.0052**	-0.0036	-0.0016
	Small	-0.0204***	-0.0204***	-0.0048	-0.0047	-0.0079**	-0.0046
	Large	-0.0184***	-0.0190***	-0.0041*	-0.0044*	-0.0016	-0.0004
	Value	-0.0216***	-0.0221***	-0.0053*	-0.0055*	-0.0049	-0.0037
	Growth	-0.0192***	-0.0194***	-0.0046**	-0.0049**	0.0002	0.0002
SMB	All	0.0056	0.0072	0.0066	0.0063	-0.0035	-0.0007
	Small	-0.0126*	-0.0113*	-0.0023	-0.0023	0.0021	-0.0049
	Large	0.0123**	0.014***	0.0078	0.0072*	0.0001	0.0016
	Value	0.0075	0.010	0.0059	0.0054	0.0020	-0.0004
	Growth	0.0088*	0.0102**	0.0054	0.0050	0.0002	0.0013
HML	All	-0.0116***	-0.0106***	0.0030	0.0034	-0.0005	-0.0017
	Small	-0.0118**	-0.0116**	0.0048	0.0047	-0.0017	-0.0034
	Large	-0.0091**	-0.0076*	0.0029	0.0036	0.0010	0.0003
	Value	-0.0222***	-0.0211***	0.0017	0.0021	0.0022	0.0013
	Growth	-0.0024	-0.0020	0.0041	0.0046	-0.0021	-0.003
NOISE	All		0.0146*		-0.0089		0.0134
	Small		0.0100		0.0042		0.0221
	Large		0.0171**		-0.0165		0.0076
	Value		0.0219**		-0.0129		0.0079
	Growth		0.0117*		-0.0105		0.0131
$\Delta R^2$	All	0.080	0.082	0.019	0.032	0.064	0.102
N	All	258,529	254,405	53,764	53,288	13,664	13,502

**Table 3- 6: Source of mispricing: ETF or NAV?**

This table reports results from estimating pooled OLS regressions of the following:

$$R_{i,t+1} = a_i + \sum_{k=0}^3 b_k PREM_{i,t-k} + \sum_{k=0}^3 c_k R_{i,t-k} + e_{i,t+1}$$

where  $R_{i,t+1}$  is the raw ETF return (in Panel A), or the NAV return (in Panel B).  $PREM_i$  is the ETF premium relative to NAV. Standard errors are adjusted for serial and cross-sectional correlation as in Driscoll and Kraay (1998). I also provide an F-test for the hypothesis that the sum  $\sum b_k$  equals zero. \*/\*\*/\*\*\*/ denotes statistical significance at the 10, 5 and 1 %.

Variables	Sample:					
	Full	Sector	Core	Large	Mid	Small
<b>PANEL A: Y = ETF return, t</b>						
PREM, t-1	-0.570* (1.86)	-0.511* (1.76)	-0.677 (1.55)	0.016 (0.03)	-0.850 (1.34)	-1.284** (2.48)
PREM, t-2	-0.964*** (3.01)	-0.860** (2.60)	-1.142*** (3.26)	-1.366** (2.53)	-1.113** (2.37)	-1.011** (2.48)
PREM, t-3	-0.352 (0.92)	-0.257 (0.80)	-0.531 (0.94)	-0.553 (0.85)	-0.897 (1.22)	-0.317 (0.59)
PREM, t-4	0.071 (0.20)	0.015 (0.05)	0.200 (0.44)	0.348 (0.53)	0.402 (0.69)	-0.076 (0.20)
ETF Ret, t-1	-0.091*** (3.23)	-0.095*** (3.50)	-0.086*** (2.68)	-0.091*** (2.75)	-0.058* (1.81)	-0.101*** (2.89)
ETF Ret, t-2	-0.046 (1.25)	-0.043 (1.25)	-0.052 (1.21)	-0.056 (1.24)	-0.053 (1.19)	-0.047 (1.13)
ETF Ret, t-3	-0.007 (0.24)	-0.009 (0.30)	-0.005 (0.15)	0.004 (0.10)	-0.011 (0.30)	-0.014 (0.40)
ETF Ret, t-4	-0.014 (0.42)	-0.016 (0.48)	-0.011 (0.30)	-0.004 (0.09)	-0.016 (0.42)	-0.016 (0.45)
F-test: $\sum_{k=0}^3 b_k = 0$	7.79***	6.91**	7.98***	3.17*	3.84*	13.20***
$R^2$	0.012	0.012	0.014	0.017	0.011	0.018
N	260,449	134,621	125,828	66,126	28,559	31,143
<b>PANEL B: Y = NAV return, t</b>						
PREM, t-1	0.358 (1.15)	0.468 (1.62)	0.147 (0.34)	0.756 (1.54)	0.216 (0.33)	-0.489 (0.93)
PREM, t-2	-1.040*** (3.25)	-0.921*** (2.80)	-1.238*** (3.42)	-1.414** (2.51)	-1.273** (2.59)	-1.102** (2.60)
PREM, t-3	-0.351 (0.87)	-0.251 (0.72)	-0.542 (0.92)	-0.551 (0.80)	-0.965 (1.26)	-0.329 (0.60)
PREM, t-4	0.043 (0.12)	-0.005 (0.02)	0.157 (0.32)	0.365 (0.51)	0.380 (0.62)	-0.183 (0.45)
NAV Ret, t-1	-0.090*** (3.13)	-0.094*** (3.39)	-0.085** (2.61)	-0.093*** (2.78)	-0.051 (1.57)	-0.099** (2.80)
NAV Ret, t-2	-0.051 (1.35)	-0.047 (1.35)	-0.057 (1.30)	-0.062 (1.33)	-0.058 (1.30)	-0.052 (1.21)
NAV Ret, t-3	-0.008 (0.28)	-0.009 (0.33)	-0.007 (0.19)	0.002 (0.04)	-0.011 (0.30)	-0.017 (0.50)
NAV Ret, t-4	-0.015 (0.44)	-0.017 (0.49)	-0.013 (0.34)	-0.004 (0.10)	-0.018 (0.45)	-0.019 (0.53)
F-test: $\sum_{k=0}^3 b_k = 0$	2.16	1.26	3.29*	0.78	1.50	7.61**
$R^2$	0.014	0.014	0.014	0.020	0.011	0.015
N	259,800	134,350	125,450	65,904	28,445	31,101

**Table 3- 7: Comovement in relative trading activity and liquidity**

Daily shocks to relative turnover (price impact) for ETF  $i$  are regressed on equally-weighted shocks in relative turnover (liquidity) of other ETFs in the own- and distant styles:

$$\Delta REL(TO)_{i,t} = \alpha_i + \sum_{j=-1}^{+1} \left( \beta_{i,j,O} \Delta REL(TO)_{OWN,t+j} + \beta_{i,j,DI} \Delta REL(TO)_{DIST,t+j} \right) + u_{i,t}^{TO}$$

$$\Delta REL(PI)_{i,t} = \alpha_i + \sum_{j=-1}^{+1} \left( \beta_{i,j,O} \Delta REL(PI)_{OWN,t+j} + \beta_{i,j,DI} \Delta REL(PI)_{DIST,t+j} \right) + u_{i,t}^{PI}$$

The regressions include concurrent, lagged and lead values for the factors. I report the average and median values for the concurrent, lagged, lead, sum coefficients and the average  $R^2$ ; the percentage of funds with positive coefficients, the percentage of funds with positive and significant coefficients, negative and significant coefficients. Test of statistical significance for the mean is based on the cross-sectional t-statistic for the average coefficient, while the significance for the median is based on a sign-test. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % level.

Statistic	Own-style betas				Distant-style betas				Avg. $R^2$
	Conc	Lag	Lead	Sum	Conc	Lag	Lead	Sum	
<b>PANEL A: Relative turnover</b>									
Mean	<b>0.616***</b> (26.51)	0.104*** (5.49)	0.089*** (5.13)	0.810*** (16.68)	<b>0.205***</b> (10.79)	-0.043*** (2.91)	-0.056*** (3.50)	0.107*** (2.85)	0.086
Median	<b>0.606***</b>	0.089***	0.075***	0.689***	<b>0.162***</b>	-0.041**	-0.032**	0.073***	0.084
% pos	<b>100.00</b>	71.95	65.85	94.51	<b>82.32</b>	40.24	37.20	57.93	
% pos & sig	<b>93.90</b>	23.78	23.17	87.20	<b>51.83</b>	6.71	7.32	44.51	
% neg & sig	<b>0.00</b>	11.59	4.88	0.61	<b>0.61</b>	18.29	15.24	20.73	
<b>PANEL B: Relative price impact</b>									
Mean	<b>0.972***</b> (42.66)	-0.006 (0.34)	-0.012 (1.03)	0.955*** (25.21)	<b>-0.082***</b> (4.38)	0.012 (1.06)	0.020* (1.78)	-0.049 (1.41)	0.281
Median	<b>0.982***</b>	0.030	-0.008	1.017***	<b>-0.106***</b>	0.012	0.021	-0.063***	0.257
% pos	<b>100.00</b>	54.88	48.17	96.34	<b>32.93</b>	53.66	56.10	46.34	
% pos & sig	<b>99.39</b>	25.61	10.98	93.29	<b>16.46</b>	24.39	22.56	35.98	
% neg & sig	<b>0.00</b>	28.66	22.56	1.83	<b>43.90</b>	15.24	15.24	42.68	

**Table 3- 8: Explaining the degree of excess return comovement**

This table reports results from regressions of the amount of return comovement ( $R_{ret,m}^2$ ) for ETF  $i$  on the following measures of correlated demand shocks: commonality in abnormal trading activity ( $R_{rTO,m-1}^2$ ), commonality in relative price impact ( $R_{rPI,m-1}^2$ ). Comovements are based on the model  $R^2$ , or the fraction of model  $R^2$  attributable to the own and distant style factors.  $R^2$  is estimated every month ( $m$ ) using daily data. Other variables included are expense ratios, share creation/redemption activity (CREATE), magnitude of average misvaluation [abs(PREM)], monthly ETF and underlying portfolio proportional quoted spreads (ETF QSPR, UND QSPR; signed to indicate liquidity), volatility of NAV returns (STD(NAV)), Hu, Pan and Wang (2013) funding liquidity (NOISE). All variables are measured at the end of the previous month except for the macro variables STD(NAV) and NOISE that are contemporaneous. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % level. Standard errors are double-clustered by fund and calendar time.

Variables	(1)	(2)	(3)	Core style		(6)	(7)
				(4)	(5)		
Exp. Ratio, t-1	-0.173*** (2.63)	-0.142** (2.20)	-0.143** (2.23)	-0.335*** (5.02)	-0.008 (0.09)	-0.315*** (5.53)	-0.154 (1.09)
CREATE, t-1	-6.901*** (5.10)	-6.595*** (4.87)	-6.425*** (4.79)	-4.453** (2.17)	-6.518*** (4.29)	-9.192*** (6.64)	-1.365 (1.33)
abs(PREM), t-1	-61.968*** (2.82)	-59.117** (2.53)	-61.518*** (2.64)	-73.541*** (2.72)	-42.255 (1.51)	1.333 (0.08)	-49.100** (2.28)
ETF QSPR, t-1	6.748*** (4.99)	5.896*** (4.48)	5.815*** (4.46)	6.113*** (4.46)	4.601** (2.05)	7.764*** (5.84)	5.530*** (3.08)
UND QSPR, t-1	-3.351 (1.63)	-3.073 (1.53)	-2.955 (1.47)	-4.305* (1.75)	-2.070 (0.84)	-9.215*** (3.48)	-4.144* (1.76)
STD(NAV), t	3.003** (2.22)	3.057** (2.18)	2.988** (2.10)	3.591* (1.92)	3.273** (2.19)	0.046 (0.03)	3.247** (2.04)
NOISE, t	1.740*** (3.97)	1.657*** (3.88)	1.662*** (3.86)	1.712*** (2.84)	1.322*** (3.31)		1.424*** (3.21)
$R_{rto,t-1}^2$	0.077*** (3.18)		0.049** (2.04)	0.050 (1.60)	0.049 (1.50)	0.031 (1.41)	0.048** (2.10)
$R_{rlit,t-1}^2$		0.095*** (6.77)	0.092*** (6.46)	0.056*** (3.06)	0.103*** (5.06)	0.097*** (6.25)	0.048*** (4.13)
<i>Dummies</i>							
Sector, style	YES	YES	YES	YES	YES	YES	YES
Calendar time	NO	NO	NO	NO	NO	YES	NO
ETF	NO	NO	NO	NO	NO	NO	YES
$R^2$	0.222	0.235	0.233	0.225	0.108	0.291	0.342
N obs	11,843	11,794	11,689	5,722	5,967	11,689	11,689

## **CHAPTER FOUR: LOCAL TRADING HABITATS AND COMMONALITY IN MISVALUATION: EVIDENCE FROM TWIN-SECURITIES**

### **4.1 INTRODUCTION**

There are more than 650 plain vanilla Exchange-Traded Funds (ETFs) tracking European equity indices with a stock market listing in Europe. Many of these funds are, however, duplicates of one another as they are either cross-listed from a different European exchange, or they track an identical benchmark index. The most popular benchmark index, Euro Stoxx 50, is tracked by 78 different ETFs. For instance, the shares of iShares Euro Stoxx 50 are traded on six different European exchanges, while Lyxor Euro Stoxx 50 is traded on five. The only fundamental difference between cross-listed ETFs (e.g. the shares of iShares Euro Stoxx 50 traded in London vs. Frankfurt), or more generally between ETFs passively tracking an identical benchmark index (e.g. the shares of iShares and Lyxor Euro Stoxx 50 traded in Italy vs. France), is that they are owned and traded by different groups of investors. Anecdotal evidence suggests that the primary reason for ETF providers to list cross-list existing funds (or establish new funds) is to support local clientele demand<sup>42</sup>. This provides a unique opportunity to investigate whether the trading by a local clientele can induce excessive return comovements. The theoretical motivation is

---

<sup>42</sup> See section 2 for examples.

based on “preferred habitat” theory of Barberis, Shleifer and Wurgler (2005). When local securities are predominantly owned by local investors, changes in their sentiment, risk-aversion, or liquidity needs can generate a common factor in the returns of securities in the habitat.

Empirically the challenge is how to measure excess comovement because there is no consensus on how to measure comovements due to common fundamentals. To control for common fundamentals, I use return differentials between twin-securities traded in different countries. The existing literature has also used the twin-security approach. Lee, Shleifer and Thaler (1991) argue that changes in the premiums of U.S. Closed-End Funds relative to NAV tend to move together due to unpredictable variations in local sentiment that are impounded in CEF prices. Bodurtha, Kim and Lee (1995) find that changes in the premiums of international country CEFs comove positively (negatively) with local (foreign) market returns. Similar evidence is also found for pairs of “Siamese twins” such as Royal Dutch and Shell (Froot and Dabora, 1999) and for cross-listed stocks (Gagnon and Karolyi, 2010).

An important contribution of this study is to provide tests that differentiate between the two possible sources of excess comovement, namely common non-fundamental demand shocks and the alternative explanation of differences in information diffusion between twin-securities (Barberis, Shleifer and Wurgler, 2005).

Differences in information diffusion can arise in three different ways. First, it can arise mechanically in the presence of non-trading when trading hours are misaligned. For example, the prices of U.S. traded securities will generally reflect local news as it arrives, while foreign markets will react the following morning. This is a concern in Bodurtha, Kim and Lee (1995) because they rely on Friday’s closing prices for both local and foreign markets, and in Froot and Dabora (1999) because they use “Siamese-twin” securities traded in the U.S. vs. Europe (five hour misalignment in trading hours). Second, stale pricing between twins can generate differences in information diffusion: more actively traded securities will comove “too much” with one another compared to securities that have stale prices. This is a concern not only for CEFs that typically invest in illiquid securities (Cherkes, Sagi and Stanton, 2008), but also for cross-listed stocks in the U.S.. While Gagnon and Karolyi (2010) use time synchronized prices whenever possible for their sample of ADRs traded in the U.S., trading activity in the middle of the U.S. day (at the European market close) is lower and stale pricing is therefore a potential concern, and for Asian ADRs it is not even possible to obtain time-synchronized prices. Third, if

local investors are better or faster at processing local information about fundamentals, then locally traded securities will be more price informative about local fundamentals relative to foreign securities. Eun and Sabherwal (2003) investigate Canadian stocks listed on both the Toronto Stock Exchange and a U.S. exchange and find that the U.S. share of price discovery is directly related to the U.S. share of trading activity and the ratio of U.S. liquidity to home liquidity, while Agarwal, Liu and Rhee (2007) demonstrate that for Hong-Kong stocks cross-listed in London, price discovery occurs mainly in the home market.

To make my tests as clean as possible, I only include funds tracking either regional (Pan-European), or individual country indices for the six ETF markets included in this study (France, Germany, Italy, Netherlands, Switzerland and the U.K.). The majority of ETFs in the final sample are regional ETFs (292 total listings) followed by country funds (124). I exclude all actively managed, fundamentally-weighted and leveraged funds leaving me with a final sample of 416 funds spanning 8 years and with assets of \$140 billion as of 11/2013.

Without limits-to-arbitrage, there can be no misvaluation between twin securities. ETFs have an open-end structure via the share recreation/redemption process that facilitates arbitrage by allowing Authorized Participants (APs) to exchange the underlying basket for ETF shares with the ETF manager when there is excess demand for the ETF, and vice-versa when the opposite occurs. However, arbitrage remains limited by direct and indirect costs. In comparison to the U.S. market where misvaluation among domestic equity ETFs is relatively low (as measured by the absolute value of ETF premium relative to NAV) and liquidity is high (see Chapter 3), European ETFs have higher misvaluation due to a fragmentation of liquidity across multiple exchanges, risk of settlement failure and costs associated with the transfer and settlement of ETF shares across national markets. Additional costs faced by cross-border arbitrageurs include the requirement to have depository accounts in each country, follow post-trading rules in each jurisdiction, and specific costs that depend on the type of ETF listing.

To preview my results, I show that twin return differentials (between an ETF and its own shares traded abroad) contain a strong country-specific common factor which I refer to as the local habitat factor. This common factor accounts for almost half of the variation in twin return differentials. Economically the commonality in twin return differentials is strong: a 1 Std. Dev. increase in the local habitat factor is associated with a 20.6 bp increase in daily return differentials. The proportional impact, measured relative to the variability in raw ETF returns, is

also considerable at 15.34 percent. The strength of this effect declines rapidly with the return horizon. At the monthly horizon, the proportional impact of a shock to the local habitat factor accounts for only 3.75 percent of the variation in raw returns. Exchange rates do not matter because sub-sample results for the Eurozone countries (with local and foreign representing locations only within the Eurozone) are very similar to the full sample results. A placebo test confirms that twin return differentials of ETFs from the same trading location (e.g. a pair of ETFs tracking FTSE 100, both traded in the UK) are not exposed to the local habitat factor.

Non-trading is generally not a concern in this study because trading hours among the European markets studied are almost perfectly aligned. Stale pricing is a potential concern that I address. First, I exclude ETFs from the sample that have excessive non-trading during the entire sample period, or individual fund-months with excessive non-trading. Next, I compare the degree of commonality in twin return differentials among ETFs with identical benchmark indices, but different levels of liquidity and non-trading probability. Commonality in return differentials actually increases as the liquidity of the ETF in the short-leg increases, holding constant the liquidity of the ETF in the long-leg. This finding is opposite to the stale pricing argument, but it is consistent with local-trading induced misvaluation when local demand shocks reflect short-term demand. This is because liquidity is more valuable to short-term investors relative to long-term (Amihud and Mendelson, 1986; also see the results in Chapter 2).

To rule out the alternative explanation of commonality in misvaluation due to differences in the speed of information diffusion, I conduct three tests. First, I focus on a sub-sample of ETFs tracking Pan-European equity indices where local investors are unlikely to have an informational advantage. The results are not materially different in this sample. Second, I design a cleaner experiment where local investors are highly unlikely to have an informational advantage. Specifically, I compare the degree of commonality in misvaluation for two sub-samples: a) the trading location of ETF-pairs is different from to the trading location of their underlying securities vs. b) the trading locations of the ETF-pairs and their underlying securities match. The results do not materially differ between the two samples. Third, I show that premiums of the local ETF (relative to NAV) predicts a reversal in future returns for the local ETF, while premiums of the foreign twin (relative to NAV) has little impact on future returns of the local ETF, which is more consistent with the reversal of local demand shocks as opposed to local ETF prices catching up to fundamental news already incorporated into foreign ETF prices.

Korniotis and Kumar (2013) argue that local misvaluation can be caused by undiversifiable incomes shocks that affect the risk-aversion or risk-sharing ability of local investors. Due to limited arbitrage, cross-border arbitrageurs cannot immediately absorb this liquidity demand, which can generate persistence in misvaluation. This mechanism could expose both locally traded stocks and ETFs to local-trading induced misvaluation. Consistent with this hypothesis, I find that twin return differentials are positively (negatively) correlated with stock market returns in the local (foreign) market. This relationship is strongest at the daily horizon, weaker but still significant at weekly horizon, and disappears entirely at the monthly horizon, which is consistent with a smaller fraction of misvaluation in local stock returns at longer horizons.

The motivation for local clienteles in ETFs is closely related to the literature on home and local bias. Home bias, or the tendency for investor to invest in equity securities at home, is pervasive around the world (e.g. Chan, Covrig and Ng, 2005), while local bias reflects the tendency for investors to invest in nearby domestic securities (e.g. Coval and Moskowitz, 2001). While there are several potential explanations for the existence of home bias<sup>43</sup>, many are not relevant here. For instance, there is no exchange-rate uncertainty for local investors trading locally traded ETFs denominated in the local currency. Also, information asymmetry between local and foreign investors is less relevant in the context of ETFs tracking foreign indices. The preference for locally traded ETFs can arise if local investors have lower deadweight costs when investing in the domestic market (or higher costs when investing abroad) relative to the world investor (Chan, Covrig and Ng, 2005). In the case of European ETFs, local investors face additional costs for cross-border trades related to the trading, transfer and settlement of ETF trades as well as legal requirements to have depository accounts in each country and follow post-trading rules in each jurisdiction. Local investors may therefore prefer to invest in locally traded ETFs in order to gain exposure to foreign stock markets.

This paper is also related to a growing literature studying the importance of ownership structure on asset prices. Prinsky and Wang (2006) document anomalous return comovements among neighbouring firms headquartered in the same geographical area. Garcia and Norli (2012) find that truly local firms (based on the locality of their business operations) earn higher returns than those of geographically dispersed firms, presumably because local firms are risky due to

---

<sup>43</sup> E.g. transaction costs, real exchange rate risks, information barriers, corporate governance issues, familiarity, loyalty, patriotism and overconfidence (see Bekeart and Wang, 2009; Van Nieuwerburgh and Veldkamp, 2009).

lower investor recognition (as in Merton, 1987). Bartram, Griffin and Ng (2014) find that foreign ownership linkages affect stock returns. My contribution in this regard is to offer a cleaner test to filter out the fundamental variation in asset returns.

The study proceeds as follows. Section 2 reviews the theories and presents testable implications. Section 3 describes the data, formally defines misvaluation and presents summary statistics. Section 4 presents empirical tests for local habitats. Section 6 provides further tests to disentangle misvaluation from information diffusion. Section 7 provides additional robustness tests and Section 7 concludes.

## 4.2 LOCAL TRADING HABITATS AND TESTABLE HYPOTHESES

Barberis, Shleifer and Wurgler (2005) (henceforth BSW) formalize a “habitat” view of return comovement where a) some investors limit their trading to a particular subset of securities, such as local stocks; or b) some securities are primarily owned by a specific subset of investors, such as local investor. The correlated trading of these investors can then generate a common factor in the returns of securities in the habitat. In BSW’s Eq. (4), asset returns are given by two factors:

$$R_{i,t} = \varepsilon_{i,t} + \gamma_{i,L} \Delta u_{L,t} \quad \text{where } i \in L \quad (17)$$

The first component reflects news about fundamental value ( $\varepsilon_t$ ), which is often characterized via common industry, country or global return factors in the international finance literature (see e.g. Bekaert, Hodrick and Zhang, 2009). The second component reflects common *demand* shocks for the investors in habitat  $L$  ( $\Delta u_{L,t}$ ), which can arise from changes in investors’ sentiment, risk aversion or liquidity needs.

In this study  $L$  denotes a local (or country “ $L$ ”-specific) trading habitat and  $i$  denotes an ETF that is traded on the local market. A local ETF habitat can arise when local investors restrict their trading to local ETFs, or more generally when local ETFs are primarily owned by local investors. The reason why local ETF habitats are likely to exist is because local ETFs provide investors with easy access to foreign stock market indices without having to incur costs associated with the trading, settlement or transfer of ETF shares across national markets, operational costs that arise from the requirement to have depository accounts in each country and follow post-trading rules in each jurisdiction. Other costs that can be avoided include withholding taxes, which apply to dividend income that foreign investors (non-residents) pay to

the government where the stocks are domiciled, and in many cases also to the government where the ETF is located<sup>44</sup>. International trading restrictions and lack of information (Merton, 1987) or constraints on investors' attention (Peng and Xiong, 2006) can also contribute to the formation of local habitats. Anecdotal evidence also suggests that ETF providers list new (or cross-list existing) funds primarily to support local clientele demand<sup>45</sup>.

#### 4.2.1 Identifying the local habitat factor and testable implications

In order to control for fundamental variation in returns, I use the return difference between ETF  $i$  and its twin security  $p_i$  (defined shortly):

$$R_{i,t} - R_{p_i,t} = (\varepsilon_{i,t} - \varepsilon_{p_i,t}) + (\gamma_i \Delta u_{L,t} - \gamma_{p_i} \Delta u_{F,t}), \text{ where } i \in L, p_i \in F \quad (18)$$

where  $R_i$  = return for ETF  $i$  or its twin  $p_i$  at time  $t$

$\Delta u_{L,t}$  = common demand shock in country  $L$  where ETF  $i$  is traded

$\Delta u_{F,t}$  = common demand shock in country  $F$  where the ETF  $p_i$  is traded

ETFs can have multiple twins. There are two approaches to form pairs of twin ETFs. First, I use an ETFs own (cross-listed) shares traded on a foreign market. For instance, the shares of iShares Euro Stoxx 50 are traded in the Netherlands, Germany and Italy (in EUR), and in London (in GBP) and Switzerland (in CHF). In this case the fundamentals of the shares, wherever they are traded, are by definition the same. However, even identical ETF shares trading in different markets are not fully fungible due to direct and indirect costs associated with the trading, settlement and transfer of ETF shares across national markets; arbitrage remains risky. Second, the twin can be chosen from among the set of foreign ETFs that passively track the same underlying index. Consider, for example, iShares Euro Stoxx 50 traded in the Netherlands. The foreign twins of this ETF include (but is not limited to) Lyxor ETF Euro Stoxx 50 (traded in France and Germany), and HSBC Euro Stoxx 50 ETF (traded in France and the U.K.). In the case of ETF pairs tracking an identical index, both ETFs hold the same underlying basket of securities (aside from differences that arise from different replication methods), implying that a

<sup>44</sup> see Appendix A3 for additional details on the mechanics and costs of cross-border trades

<sup>45</sup> Examples of press releases by: State Street Global Advisors (28.01.14), HSBC Global Asset Management (27 September, 2013), UBS (20 March 2013), Source UK Services Ltd (11 September 2014)

long-short position can be established with little fundamental risk<sup>46</sup>. Table 4-1 provides additional examples and further information on selecting twins.

In either case ETF  $i$  and its twin  $p_i$  are both claims to the same underlying securities and they should therefore inherit the systematic risks of their underlying securities (facilitated by arbitrage). Hence, the twin return differential effectively controls for news about fundamental value ( $\varepsilon_{i,t} - \varepsilon_{p_i,t} = 0$ ). However, ETF  $i$  and  $p_i$  are traded in different markets ( $L$  and  $F$ ) and can be influenced by correlated local demand shocks ( $\Delta u_{L,t} - \Delta u_{F,t}$ ). If trading location matters for misvaluation, then we should expect to find a common factor in the twin return differentials of ETF pairs traded in countries  $L$  and  $F$ .

**Hypothesis 1:** *Twin return differentials are correlated across ETF-pairs traded in the same set of countries ( $L$  and  $F$ ). Specifically,  $\text{Corr}(R_i - R_{p_i}, R_j - R_{p_j}) > 0$ , where  $i, j \in L$  and  $p_i, p_j \in F$ .*

Such excess comovements can, according to the behavioural view, arise as a result of noise trader sentiment (DeLong *et al.*, 1990; Lee, Shleifer and Thaler, 1991). In this case the term  $\Delta u_{L,t} - \Delta u_{F,t}$  can be interpreted as the differential sentiment shock between local investors in country  $L$  and  $F$ . Alternatively, local demand shocks can also arise in a rational, but frictions-based setting. Kumar and Korniotis (2013) argue that, when levels of local ownership are high, changes in the risk-aversion of local investors and changes in their ability to engage in risk sharing can generate misvaluation among local securities. To support their argument, the authors show that a geography-based trading strategy earns significant abnormal returns and that the predictability in local returns is arbitrated away by non-local investors within a time-period of one year. When local stock ownership is high, we should expect to find a positive correlation between twin return differentials  $R_i - R_{p_i}$  and local stock market returns since both are exposed to correlated local trading in countries  $L$  and  $F$ :

**Hypothesis 2:** *The return differential between ETF  $i$  (traded in country  $L$ ) and its foreign peer  $p_i$  (traded in country  $F$ ) is positively correlated with the stock market returns in country  $L$  and negatively correlated with the stock market return in country  $F$ .*

---

<sup>46</sup> An alternative arbitrage strategy is to use the share creation mechanism. Consider two ETFs A and B traded in France and Germany respectively. Suppose ETF A is trading at \$150, while ETF B is trading at \$75 with the fundamental price at \$100. An arbitrage trade would then involve buying the underpriced ETF (B), redeem it for underlying securities using the share creation/redemption mechanism, create new ETF B shares using the underlying securities that were just received and sell the over-priced ETF B shares in Germany.

An alternative explanation to the demand-based view of excess comovements is differences in the speed of information diffusion between twin securities traded in different countries (Barberis, Shleifer and Wurgler, 2005). In this case ETF returns are given by:

$$\begin{aligned} R_{i,t} &= \varepsilon_{i,t} = e_{i,t} + e_{L,t} + (1-\mu)e_{F,t} \quad \text{where } i \in L \\ R_{p_i,t} &= \varepsilon_{p_i,t} = e_{p_i,t} + (1-\mu)e_{L,t} + e_{F,t} \quad \text{where } p_i \in F \end{aligned} \tag{19}$$

where  $\mu$  = fraction of news incorporated with a lag

In Eq. (3) the returns of locally traded securities reflect local news instantaneously, while foreign news is incorporated with a lag. Taking the return differential of ETF  $i$  and  $p_i$  would only eliminate the news about fundamental value unrelated to country  $L$  and  $F$  ( $e_{i,t} - e_{p_i,t} = 0$ ). This yields the same predictions as in Hypothesis 1 and 2.

Such differences in information diffusion can arise mechanically due to systematic differences in non-trading (or stale pricing) between local and foreign twin ETFs. For example, if the shares of iShares Euro Stoxx 50 traded on the Italian stock market are less liquid and more prone to non-trading (or stale pricing) relative to the ETFs shares traded in France, then it would mechanically create a divergence in return differentials that is positively (negatively) correlated with local (foreign) stock market returns, and positively correlated with the twin return differential of other ETF pairs traded in the same country pairs that also suffer from stale pricing. In order to ensure that non-trading is not driving the results, I eliminate ETFs with excessive non-trading (more details in the next section) and perhaps even more importantly, I analyze differences in the degree of commonality in twin return differentials for sub-samples of ETF-pairs sorted on their liquidity characteristics.

There is another, more interesting, story based on differences in information diffusion. Specifically, if local investors are better or faster at processing information about abstract local fundamental factors, then it can give rise to commonality in the return differential between local ETFs relative to their foreign twins. Van Niewerburgh and Veldkamp (2009) present a model where investors have a small home information advantage. Information asymmetry persists in their model not because investors cannot learn what locals know, nor because such information is expensive, but because investors choose not to learn. It is simply more profitable to specialize. Recent studies have also confirmed that local investors have an informational advantage over non-local investors: local analysts make more precise earnings forecasts (Bae, Stulz and Tan,

2008), hedge funds with a local research office presence outperform other hedge funds (Teo, 2009), the institutional ownership of geographically proximate investors is a better predictor of future returns relative to the ownership of non-local investors (Baik, Kang and Kim, 2010). To rule out this alternative explanation, I perform tests based on: 1) reversals in the degree of commonality in twin return differentials over longer horizons and 2) a comparison of the degree of commonality in twin return differentials for sub-samples where the ETF owners differ in their likelihood for having an informational advantage (e.g. ETFs traded in countries other than the trading location of the ETFs underlying securities vs. ETFs traded in the same countries as their underlying securities).

#### 4.2.2 Key variables

The main dependent variable in this study is the twin return differential,  $R_{i-p,t} \equiv R_{i,t} - R_{p,t}$ . In order to more accurately capture the impact of local-trading induced misvaluation, we need to remove the impact of arbitrage trades on the twin return differentials, which can be accomplished by a regression of the twin return differential ( $R_{i-p,t}$ ) on the lagged twin price differential ( $P_{i-p,t-1}$ ). Before discussing the rationale any further, I first need to define the price differential. The standard definition of ETF misvaluation is the premium, or the log differential between the market price of an ETF and the market value of the ETF's portfolio on a per-share basis ( $NAV$ <sup>47</sup>):

$$PREM_{i,t} = \ln(P_{i,t} / NAV_{i,t}) \quad (20)$$

where:  $P_{i,t}$  = ETF mid-quote price at the close of day  $t$  for ETF  $i$ ,  
 $NAV_{i,t}$  = Net Asset Value per share for ETF  $i$  on day  $t$

To mitigate the possibility that prices are affected by microstructures issues such as bid-ask bounce, I use closing mid-quote prices rather than end-of-day prices. Engle and Sarkar (2006) also suggest using mid-quote prices in order to mitigate concerns about the illiquidity of the shares of smaller ETFs. Next, I define two versions of the twin price differential:

$$\begin{aligned} P_{i-p1,t} &= \ln(P_{i,t} / P_{p1,t}) \\ P_{i-p2,t} &= \ln(P_{i,t} / P_{NAV_{i,t}}) - \ln(P_{p2,t} / NAV_{p2,t}) = PREM_{i,t} - PREM_{p2,t} \end{aligned} \quad (21)$$

where  $P_{p1,t}$  = the price of an ETF's (own) shares traded on a foreign market,

---

<sup>47</sup> NAV also includes accrued income from securities lending, underlying stock dividends and cash.

$P_{p2,t}$  = the price of a different ETF with an identical benchmark index

The first,  $P_{i-p1}$ , is the log price differential between ETF  $i$  and its shares traded abroad. The share price of an ETF should not depend on its trading location (in efficient markets).  $P_{i-p2}$  is the log price differential between two different ETFs, ETF  $i$  and a foreign traded ETF with the same benchmark index. Because price levels are not directly comparable across different ETFs, I scale the ETF prices by their corresponding NAV prices before taking the price difference. This corresponds to taking the difference in premiums (relative to NAV) between ETF  $i$  and its twin. The interpretation is straightforward. Consider an ETF tracking FTSE 100. If the shares of the ETF are trading at a premium (relative to NAV) of 20 basis points, while the ETFs foreign peer tracking FTSE 100 trade at a discount of 30 basis points, then the former is overpriced relative to the latter by a magnitude of 50 basis points. Petajisto (2013) use a similar method to calculate the misvaluation of U.S. traded ETFs that track international equities (and other illiquid asset classes) in an effort to control for stale pricing in the underlying securities.

The twin return differential  $R_{i-p,t}$  is equivalent to the change in  $P_{i-p,t}$  when log-returns and prices are used<sup>48</sup>. To keep this link clear, I will use log-returns throughout this study.

Even if  $R_{i-p,t}$  perfectly controls for fundamental risk (as in Eq. (2)), then the return differential reflects not only the impact of local demand shocks, but also the subsequent arbitrage trades and the associated reversal in misvaluation. Suppose that local investors buy ETF  $i$  in the local market causing price pressure, which results in a price appreciation of ETF  $i$  relative to its foreign twin  $p_i$  ( $P_{i-p,t} > 0$ ). In this case arbitrageurs can buy the relatively underpriced foreign ETF and sell the relatively overpriced local ETF, and wait for prices to converge. This will eventually lead to a reversal in twin return differentials ( $R_{i-p,t} < 0$ ). This mechanism generates mean-reversion in changes in misvaluation. Given that demand shocks are theoretically expected to be uncorrelated (or positively auto-correlated due to some persistence in demand shocks), we can remove the effect of arbitrage trades by regressing  $R_{i-p,t}$  on  $P_{i-p,t-1}$ .

---

<sup>48</sup> This definition is exact for shares of the same ETF traded in different markets ( $\Delta P_{i-p1,t} \equiv R_{i-p1,t}$ ). However, when an ETF is compared against a different ETF with an identical benchmark, we obtain: ( $\Delta P_{i-p2,t} = R_{i-p2,t} + \Delta \ln NAV_i - \Delta \ln NAV_{p2}$ ). The last two terms drop out when ETF  $i$  and its twin  $p_i$  have identical NAV returns (i.e. when both ETFs use identical replication methods).

### 4.3 DATA

My data selection starts with all ETFs traded on the main exchanges in France, Germany, Italy, Netherlands, Switzerland and the U.K. I keep funds that i) invest in any of the six aforementioned countries or in the regional (Pan-European) market, ii) are passively managed, iii) have reached a minimum size of \$15 million<sup>49</sup> and iv) have at least 2.5 years of complete data available. To mitigate the possibility of stale pricing induced comovement in twin return differentials, I also discard v) any funds from the sample with more than 20 percent non-trading days over the full sample, or vi) individual fund-month observations where the non-trading days are in excess of 25 percent. These exclusion criteria do not have a major impact on the size of the sample<sup>50</sup>. I focus on a recent sample period from January 2006 to December 2013 when the size of the entire market is reasonably large and the data quality is good.

My main data source for the sample of ETFs, ETF prices, NAV prices and underlying indices is Bloomberg. I collect data on the number of shares outstanding and Assets Under Management (AUM) primarily from Morningstar Direct instead of Bloomberg because Morningstar has fewer misplaced decimals and higher accuracy (Bloomberg shares are rounded to the nearest thousand units). In order to verify the accuracy of the NAV data, I collect two sets of NAV prices. The first is the official NAV provided by the fund provider (available from both Bloomberg and Morningstar). The second is the closing price for the intraday NAV (iNAV). iNAV is disseminated every 15 seconds throughout the trading day by a third party data provider who has been selected by the fund sponsor (from Bloomberg). Since I use closing iNAV prices the calculation is done at the same time as for NAV prices, but by different entities.

I screen the ETF and NAV price data for a number of obvious data errors. These include screens for misreported currency denominations, stale NAV prices when the underlying market has traded and the occasional misplaced decimal. A detailed treatment of these cases is discussed in Appendix A1.

Table 4-2 provides a snapshot of the number of ETF listings as well as the Assets Under Management. The number of listings is broken down to two categories: original or primary

---

<sup>49</sup> I exclude the first 6 months of a fund's history since the data is often extreme, leaving me with an estimation sample of at least 2 years. For instance, the liquidity of newly created ETFs can be low when they are first created.

<sup>50</sup> According to Blackrock's ETF Global handbook from 09/2012, the total assets under management (AUM) of European single country funds and pan European funds are \$61.3 billion and \$53.3 billion respectively. This compares well with my sample that has a total AUM of \$55.9 billion and \$45.8 billion respectively.

listings (where the original IPO took place) and cross-listings (subsequent to the initial IPO). At the end of 2013, the total number of European equity ETF listings (main) is 416 (182). At the start of my sample (01/2006), the total AUM is \$22.6 billion. Subsequently the total AUM grows to \$86.1 billion in 01/2010, and to \$140.1 billion at the end of the sample (12/2013). Overall, there are more than twice as many ETFs that track regional as opposed to country indices (292 vs 124), but regional funds are generally much smaller. The bulk of regional ETFs (both number of and in terms of AUM) track European Large-Cap indices. Germany has the largest number of both total and primary ETF listings (126 and 79), while nearly all funds traded in Italy and the Netherlands are secondary-listings (70/71 and 16/18). In terms of the domicile of ETFs underlying securities, the most popular location in terms of number of listings is U.K. (52 total listings), while ETFs tracking German country indices have the biggest AUM (\$31.8 billion). The share of synthetically replicated ETFs (i.e. funds that use swaps to gain exposure to an index) has declined over time; the share was 42 % in 01/2006, 38 % in 01/2010 and 30 % in 11/2013.

[Table 4-2]

#### 4.3.1 Descriptive statistics

Table 4-3 provides summary statistics for the various proxies for the level and changes in twin price differentials, at the weekly horizon. First, the level and change in premiums is zero on average, and at the median, which suggests that across funds and time ETFs are overall efficiently priced. However, there is considerable variation around the mean as indicated by the standard deviation of 0.28 % and 0.65 % for the level and change in premiums. The extreme right and left tails (1 and 99 percentiles) are roughly +/- 80 bps for levels of premiums, and +/- 2 % for changes in premiums. The variability of ETF premiums is also considerable when compared against the variability in raw ETF returns: the ratio  $STD(\Delta PREM_{i,t})/STD(R_{i,t})$  is roughly 45 %.

I also make a comparison with the magnitude of premiums of U.S. equity funds in Chapter 2. In his study, the volatility of weekly premium levels and changes are 0.127 % and 0.39 % respectively. Restricting the sample to large-cap non-sector funds (to retain comparability with my European ETF sample) further reduces the volatility of the level (change) of premiums to

0.095 % (0.29 %) <sup>51</sup>. Overall, the indication is that European equity ETFs have considerably higher variations in premiums when compared to domestic U.S. ETFs, presumably due to higher arbitrage costs.

The summary statistics for the twin return differential based on an ETF and its own foreign traded shares ( $R_{i-p1,t}$ ) indicate an even higher variability of 0.96 % per week, despite the tight arbitrage link between an ETFs various listings across Europe. The average ETF in the sample has roughly 3.6 cross-listings implying that for every ETF,  $R_{i-p1,t}$  can be calculated in 3.6 different ways. In computing the twin return differential of ETFs with identical benchmark indices ( $R_{i-p2,t}$ ), the average (median) number of pairs is as high as 94 (18). Overall, the variability of  $R_{i-p2,t}$  is comparable to the variability of  $R_{i-p1,t}$  indicating that the arbitrage link between an ETFs own shares is no more efficient than between an ETF and its foreign peers tracking the same index. In total there are 400 different pairs of ETFs using an ETFs own foreign traded shares, and as many as 2727 different ETF pairs with identical benchmarks.

[Table 4-3]

#### 4.4 EMPIRICAL TESTS OF LOCAL HABITATS

Hypothesis 1 predicts a positive correlation between twin return differentials  $R_{i-p,t}$  and  $R_{j-p,t}$  for ETF-pairs traded in the same country-pairs ( $i, j \in L$  and  $p_i, p_j \in F$ ) due to local trading induced misvaluation. Specifically, the long-leg of each ETF pair ( $i$  and  $j$ ) is exposed to local demand shocks in country  $L$  ( $\Delta u_{L,t}$ ) and the short-leg of each ETF pair ( $p_i$  and  $p_j$ ) is exposed to  $\Delta u_{F,t}$ . In order to obtain a convenient summary measure for this common factor, I use the equally-weighted twin return differential of all ETFs traded in country  $L$  relative to  $F$  (same country-pairs as ETF  $i$  and its peer  $p_i$ ):

$$R_{L-F,t} = \frac{1}{N_{twin}} \sum_{j=1}^{N_{twin}} (R_{j-p,t}) = \frac{1}{N_{twin}} \sum_{j=1}^{N_{twin}} (R_{j,t} - R_{p_j,t}), \text{ where } j \in L, p_j \in F \quad (22)$$

where  $L$  and  $F$  = countries  $L$  (“Local”) and  $F$  (“Foreign”) respectively

I compute two different versions of the factor  $R_{L-F}$ . The first ( $R_{L-F1}$ ), restricts the set of twins ( $p_j$ ) to ETF  $j$ 's own shares traded abroad, while the second ( $R_{L-F2}$ ) uses the entire set of foreign twins with identical benchmark indices. To avoid inducing a spurious correlation between  $R_{i-p}$  and the

---

<sup>51</sup> These numbers do not appear in Chapter 3, but they are based on the same data.

factors, I exclude ETF  $i$  and  $p_i$  from  $R_{L-F1}$  and  $R_{L-F2}$ . Moreover, I also exclude any other ETFs tracking the same benchmark index from  $R_{L-F2}$ . With these exclusions the factor is specific to ETF  $i$  and  $p_i$ , but to avoid unnecessary clutter, I do not add any additional subscripts to the factor. I refer to  $R_{L-F}$  as the composite local habitat factor because it is a summary measure for the relative local habitat factor  $\Delta u_{L,t} - \Delta u_{F,t}$ .

Hypothesis 1 predicts that twin return differentials of ETF-pairs traded in countries  $L$  and  $F$  are positively correlated with the equally-weighted twin return differential of other ETF-pairs traded in countries  $L$  and  $F$  ( $\beta_{L-F,i} > 0$ ).

$$R_{i-p,t} = \alpha_{0,i} + \alpha_{1,i}P_{i-p,t-1} + \beta_{L-F,i}R_{L-F,t} + \varepsilon_{i,t} \quad (23)$$

The lagged twin price differential  $P_{i-p,t-1}$  is included to account for the mean-reversion induced by arbitrage trades (see section 2.2). Regression (7) is estimated separately for twin return differentials 1) between an ETF and its own shares traded abroad ( $R_{i-p1,t}; P_{i-p1,t-1}, R_{L-F1,t}$ ) and 2) between an ETF and a different ETF traded abroad with an identical benchmark index ( $R_{i-p2,t}; P_{i-p2,t-1}, R_{L-F2,t}$ ).

I run time-series regressions of (7) for each ETF-pair using all available observations and report the mean of the estimated coefficients across ETF-pairs. In calculating the standard error for the mean coefficient, I take into account the cross-equation correlations in the estimated coefficients following Hameed, Kang and Viswanathan (2010):

$$STD(\bar{\beta}) = STD\left(\frac{1}{N} \sum_{i=1}^N \beta_i\right) = \frac{1}{N} \sqrt{\sum_{i=1}^N VAR(\beta_i) + \sum_{i=1}^N \sum_{j=1, j \neq i}^N \rho_{i,j} \sqrt{VAR(\beta_i)VAR(\beta_j)}} \quad (24)$$

where  $\sqrt{VAR(\beta_i)}$  = the White standard error of the coefficient  
 $\rho_{i,j}$  = the empirical correlation between the regression residuals for  
ETF  $i$  and  $j$ .

#### 4.4.1 Results: composite local habitat factor

Table 4-4 provides the results for estimating (7) at the daily and monthly return horizons. Panel A provides the results for the return differential between an ETF and its own own foreign shares traded abroad ( $R_{i-p1,t}$ ). The  $\beta_{L-F}$  coefficient is on average positive and highly significant. To illustrate the economic magnitude, the impact of a 1 Std. Dev. shock to the local habitat factor ( $\beta_{L-F} * Std(R_{L-F1,t})$ ) is associated with an increase in twin return differentials of 20.6 bps. The

proportional impact,  $\beta_{L-F} * STD(R_{L-F1,t})/STD(R_{i-p1,t})$  is also considerable at 41.1 % indicating that (loosely speaking) almost half of the variation in twin return differentials is driven by the local habitat factor. The distribution of the  $\beta_{L-F}$  coefficients is almost symmetric. Moreover, the coefficients are positive and individually significant already at the 5<sup>th</sup> percentile of the distribution.

The results remain similar, or even slightly stronger at the monthly horizon with the average impact of a 1 Std. Dev. shock corresponding to a 27.2 bps (or 43.7 % of the Std. Dev) increase in the dependent variable. Although the impacts remain similar to those observed in daily data, the economic magnitude is considerably smaller when compared against the variability in raw ETF returns. Specifically,  $\beta_{L-F} * STD(R_{L-F1,t})/STD(R_{i,t})$  equals 15.34 percent at the daily horizon, but only 3.75 % at the monthly horizon. These findings indicate that while the proportion of common vs. idiosyncratic misvaluation remains similar across horizons, the importance of common misvaluation relative to raw returns decreases rapidly at longer horizons consistent with arbitrage forces playing a role.

[Table 4-4]

Panel B provides the results for  $R_{i-p2,t}$ , or ETF returns relative to foreign traded ETFs with an identical benchmark index. The advantage of using foreign peers is the increase in degrees of freedom (there are 2727 ETF-pairs vs. 400 ETF-pairs when an ETFs own shares are used). Perhaps more importantly, there could be systematic differences between the liquidity of an ETFs shares in the primary and secondary (cross-listing) market. Although practitioners often claim that liquidity is similar across an ETFs various listings<sup>52</sup>, in unreported tests I find that ETFs with a main listing have significantly lower spreads and less volatile premiums relative to cross-listed ETFs. Even when twin return differentials are calculated relative to foreign ETFs with identical benchmark indices, the results remain very similar.

Exchange rates cannot explain the commonality in twin return differentials at the country-level because similar results are obtained even when I restrict the sample (and the construction of the local habitat factors) to the Eurozone countries. These results are unreported for conciseness.

Finally, the twin price differentials  $P_{i-p1,t-1}$  enter with the correct (negative) sign consistent with mean-reversion in misvaluation following arbitrage trades. The coefficient on

---

<sup>52</sup> <http://www.professionaladviser.com/etfm/feature/1866073/building-liquidity>

$P_{i-p1,t-1}$  should theoretically equal -1 if shocks to premiums fully revert over one period, instead there is a considerable degree of heterogeneity in the coefficient estimates with the 5<sup>th</sup> and 95<sup>th</sup> percentiles at -0.24 and -0.92 respectively at the daily horizon. This suggests that, in the cross-section of ETF-pairs, roughly 24 to 92 % of the misvaluation reverts over one day.

#### 4.4.2 Separating the local habitat factors for $L$ and $F$

The results in the previous section indicate that twin return differentials are positively correlated with the composite local habitat factor. However, this factor is identified only relative to ETF  $i$ 's foreign peer. In order to separately identify the local habitat factors of country  $L$  and  $F$ , and to examine whether there is any remaining comovement in twin return differentials not captured by these two factors, I also consider an alternative method. Here I use the return differential between an ETF and its underlying portfolio NAV ( $R_{i-N,t} \equiv R_{i,t} - \Delta \ln NAV_{i,t} = R_{i,t} - R_{N,t}$ ). The local habitat factors can then be calculated as follows:

$$\begin{aligned}
 R_{Local,t} &= \frac{1}{N_J} \sum_{j=1}^{N_J} (R_{j-N,t}) = \frac{1}{N_J} \sum_{j=1}^{N_J} (R_{j,t} - R_{N_{j,t}}) \quad \text{where } j \in L \\
 R_{Foreign,t} &= \frac{1}{N_K} \sum_{k=1}^{N_K} (R_{k-N,t}) = \frac{1}{N_K} \sum_{k=1}^{N_K} (R_{k,t} - R_{N_{k,t}}) \quad \text{where } k \in F
 \end{aligned} \tag{25}$$

An ETF represents a claim on its underlying basket of securities (facilitated by the share creation mechanism). In order to identify the local habitat factor from the difference between ETF and NAV returns, the ETF has to be more exposed to the local habitat factor relative to the underlying basket. This is likely to be the case because ETFs are mainly held by local investors, while their underlying securities are owned by both local and foreign investors. To obtain an even cleaner measure of the local habitat factors, I restrict the sample to ETFs tracking Pan European equity indices when calculating Eq. (9). Even with a local bias of 100 %, the proportion of assets in the underlying portfolio held by local investors is at most the proportion of market cap of the underlying basket invested in local securities. Finally, I also include a remainder factor  $R_{REST}$  to test whether there is any remaining comovement in twin return differentials not captured by the country-specific habitat factors.  $R_{REST}$  is calculated as the equally-weighted ETF-NAV return difference for ETFs traded in countries other than  $L$  and  $F$ .

The following time-series regression is estimated for each fund:

$$R_{i-p,t} = \alpha_{0,i} + \alpha_{1,i}P_{i-p,t-1} + \beta_{L,i}R_{Local,t} + \beta_{F,i}R_{Foreign,t} + \beta_{R,i}R_{REST,t} + \varepsilon_{i-p,t} \quad (26)$$

Hypothesis 1 predicts that twin return differentials are positively correlated with the local habitat factor ( $\beta_{L,i} > 0$ ), vice-versa for the foreign habitat factor ( $\beta_{F,i} < 0$ ) and no comovement with the average ETF-NAV return of ETFs outside the local and foreign markets ( $\beta_R = 0$ ).

[Table 4-5]

The results in Table 4-5 are in agreement with Hypothesis 1: the mean  $\beta_L$  and  $\beta_F$  coefficients are significantly positive and negative respectively. Economically, the impact of a 1 Std. Dev. shock to the local habitat factors of country  $L$  ( $F$ ) is associated with an increase (decrease) in ETF returns relative to its own foreign shares  $R_{i-p1,t}$  of 11.1 bps (-16.0 bps), or 23.1 % (-33.0 %) of the Std. Dev. of the dependent variable. Commonality in twin return differentials is entirely country-specific as indicated by the insignificant mean coefficient on the remainder term ( $\beta_R$ ). The results are very similar for ETF returns relative to foreign ETFs with an identical benchmark index ( $R_{i-p2,t}$ ). Overall, these findings provide further evidence to support my conjecture that local demand shocks induce commonality in twin return differentials.

#### 4.4.3 Placebo test

To provide further evidence that the local habitat factor is country-specific, I construct a placebo test using the twin return differential between an ETF and its domestic peer,  $R_{i-d,t}$ . A domestic peer refers to another ETF that is traded in the same country as  $i$  (i.e.  $i, d \in L$ ). When domestic ETFs are used both for the long- and the short-leg,  $R_{i-d,t}$  should not have any exposure to the local habitat factor in the following regression:

$$R_{i-d,t} = \alpha_{0,i} + \alpha_{1,i}P_{i-d,t-1} + \beta_{L,i}R_{Local,t} + \beta_{R,i}R_{REST,t} + \varepsilon_{i-d,t} \quad (27)$$

Eq. (11) only includes the local habitat factor for country  $L$  (there is no foreign market  $F$  in this case) and the remainder factor to capture any remaining comovement in twin returns beyond country  $L$ . The twin log-price differential  $P_{i-d,t}$  is calculated as  $P_{i-p2,t}$  in Eq. (5). The results in Table 4-6 confirm that  $R_{i-d,t}$  is on average uncorrelated both with the local habitat factor and the remainder term confirming that the local habitat factor is country-specific.

[Table 4-6]

## 4.5 LOCAL MISVALUATION IN THE STOCK MARKET

Korniotis and Kumar (2013) argue that local misvaluation might be caused by undiversifiable incomes shocks that affect the risk-aversion and risk-sharing ability of local investors. Specifically, when the local economy enters a recession the risk aversion of local investors is likely to increase while their risk-sharing ability might deteriorate. Consequently, local investors should sell some of their local securities to reduce portfolio risk, which in turn will depress current prices. Due to limits-to-arbitrage, it may take some time for cross-border arbitrageurs to fully absorb this liquidity demand, which can generate persistence in misvaluation. This mechanism could expose both locally traded stocks and ETFs to misvaluation. Consequently, we should find a positive correlation between twin return differentials and local stock returns.

Such a correlation might also reflect fundamental differences in information diffusion between local and foreign-traded ETFs. For instance, if local investors are better informed about the local economy, they might impound fundamental news about the local economy faster into the prices of local securities (including ETFs) relative to foreign-traded securities which are primarily owned by foreign investors. The returns of local ETFs would then comove excessively with other locally traded securities (excessive because in a frictionless economy all information is immediately reflected in prices). To disentangle local-trading induced misvaluation from differences in information diffusion, I investigate the relationship between twin return differentials and local stock market returns over different horizons. The former hypothesis predicts stronger comovements at higher frequencies where misvaluation is more pronounced. In contrast, information diffusion about abstract fundamental risk factors (e.g. market risk, or other risk factors such as size and *BE/ME*) may persist for longer.

It is also possible that twin return differentials are not only correlated with local stock market returns, but also with local versions of commonly used risk-factors such as the Fama and French *SMB* and *HML* factors because these factors also correlate with local-trading induced misvaluation. Specifically, smaller stocks and growth stocks are more likely to be held by local investors (Kumar and Kornioitis, 2013).

Another possible source of local risk is liquidity. Bernile *et al.* (2013) find that local (U.S. state-level) macro-economic conditions can predict the liquidity of local firms, which suggests that there is a common factor in the liquidity of local firms. The authors argue that when a large proportion of the ownership and trading of stocks is local, changes in local economic conditions

can influence the risk-aversion of local investors, which in turn has theoretically been shown to affect asset liquidity (Vayanos and Wang, 2012). Commonality in local liquidity can also affect the liquidity risk of local securities (e.g. Acharya and Pedersen, 2005). To investigate this possibility, I examine whether twin return differentials are exposed to innovations in market-wide liquidity based on the Pastor and Stambaugh (2003) measure.

#### 4.5.1 Constructing local stock market factors

I construct local versions of market returns, *SMB* and *HML* returns using price and shares outstanding data primarily from Bloomberg, supplemented by accounting data from Datastream and Worldscope similar to Fama and French (2012). To build a reliable sample, I follow the static and dynamic screens suggested by Ince and Porter (2006) and extended by Schmidt *et al.* (2014), among others. The static screens restrict the sample to common equity that are traded on the country's major exchange(s), excluding stocks with special features such as ADRs, REITs, preferred stocks, warrants, closed-end funds, exchange-traded funds etc. Several dynamic screens are used to handle obvious data errors (see Appendix A2 for additional details).

Fama and French (1993) introduce common risk factors related to size (*SMB*) and book-to-market (*HML*), while Fama and French (2012) provide an international extension with regional and global factors. In constructing the country-specific *SMB* and *HML* factors, I follow Fama and French (2012) and Schmidt *et al.* (2014).

Firm characteristics used in the construction of the risk-factors include size and book-to-market ratio (*B/M*). Size is the product of Bloomberg last trade price (*PX\_LAST*) and number of shares outstanding (*BS\_SH\_OUT*). Big stocks are those in the top 90 % of market cap, and small stocks are those in the bottom 10 %. *B/M* is the ratio of book equity of the previous fiscal year (Datastream: WC03501) to size at the same point in time. At the end of June of year  $t$ , the *B/M* breakpoints are calculated at the 30<sup>th</sup> and 70<sup>th</sup> percentiles of *B/M* for the big stocks. The *SMB* and *HML* factors are then calculated as in Fama and French (2012), with the exception that the factors are country-specific, as opposed to region (EU) specific.

I also construct local versions of the Pastor and Stambaugh (2003) liquidity measure, which is based on the idea that large order flows typically reverse for illiquid stocks. Greater illiquidity-driven reversals should be associated with a more *negative*  $\gamma$ -coefficient in the following regression:

$$R_{i,t} - R_{MKT,t} = a_i + b_i R_{i,t-1} + \gamma \text{Sign}[R_{i,t-1} - R_{MKT,t-1}] DVOL_{i,t-1} + e_{i,t} \quad (28)$$

where  $R_{i,t}$  = return of stock  $i$  on day  $t$

$R_{MKT,t}$  = value-weighted stock return for country  $L$  ( $i \in L$ ) on day  $t$

$DVOL_{i,t}$  = dollar volume in \$millions

Order flow is captured here simply by the dollar volume signed by excess stock returns. Regression (12) is estimated for each stock using daily data within a month (minimum 15 observations). To mitigate the impact of outliers, I winsorize the  $\gamma$ -estimates at the 5<sup>th</sup> and 95<sup>th</sup> percentile of the cross-sectional distribution within a country. Market-wide innovations in liquidity are based on the equally-weighted change in  $\gamma$  across all stocks in a given month and country. This metric is then filtered through an autoregressive model to eliminate any residual autocorrelation (Pastor and Stambaugh, 2003).

#### 4.5.2 Results: misvaluation and local stock market returns

Table 4-7 provides the results from regressions of twin return differentials of an ETF relative to its own shares traded abroad ( $R_{i-p1,t}$ ) on local stock market returns (specification 1), plus the local *SMB* and *HML* factors (specification 2) and the local liquidity risk factor (specification 3). Note that each local factor appears twice, once for country  $L$  and once for country  $F$ . All regressions include the lagged twin price differential  $P_{i-p,t-1}$  to account for the mean-reversion induced by arbitrage trades (see section 2.2).

As hypothesized, there is a positive and significant relationship between  $R_{i-p1,t}$  and local stock market returns of country  $L$ , and vice-versa for country  $F$ . This relationship is strongest at the daily horizon, weaker but still significant at weekly horizon, and disappears entirely at the monthly horizon. This rapid decline in strength is consistent with local-trading induced misvaluation. The reason is that the dependent variable is a measure of misvaluation, while the fraction of local stock returns attributable to misvaluation is likely to decline with the return horizon due to arbitrage. Similar results are obtained by Froot and Dabora (1999), but as noted by the authors, their findings might also be explained by less misalignment in trading hours over longer horizons.

There is also some evidence to suggest that twin return differentials are positively correlated with the local *SMB* factor, and negatively with the *HML* factor, which is consistent

with notion that local-trading induced misvaluation is more prevalent among smaller stocks and growth stocks. This effect is, however, significant only at the weekly horizon.

Panel B provides the results for  $R_{i-p1,t}$  averaged to the country-pair level. This aggregation corresponds to using the local habitat factor ( $R_{L-F1,t}$  in Eq. (6)) as the dependent variable. The goal is to demonstrate the overall strength of the relationship between local stock market returns and the local habitat factor: a decomposition of R-squared reveals that roughly 10 % of the variation in the local habitat factor can be attributed to the local stock market returns of countries  $L$  and  $F$  at the daily horizon, and 8.67 % percent at the weekly horizon (for additional details on the decomposition, see section 6.1).

[Table 4-7]

As for local liquidity risk, I do not find any significant relationship with  $R_{i-p1,t}$  at the monthly horizon. Although commonality in local liquidity does not seem to matter, the liquidity and arbitrage costs of ETF-pairs are important determinants for the strength of the commonality in twin return differentials, as I will show in the next section.

#### **4.5.3 Robustness: regional risk factors**

Previous results suggest that twin return differentials  $R_{i-p1,t}$  are positively correlated with local stock market returns and negatively correlated with foreign stock market returns. Exposure to regional or global risk factors cannot explain these findings; if some ETFs are overexposed to regional/global risk (relative to their foreign peer), others must be underexposed by definition. The result in Table 4-5 that  $R_{i-p1,t}$  are correlated only with the local habitat factors for countries  $L$  and  $F$ , but not the remainder factor which captures misvaluation in countries other than  $L$  or  $F$  is consistent with the idea that regional or global factors cannot explain the commonality in  $R_{i-p1,t}$ . Nevertheless, in unreported results I verify that  $R_{i-p1,t}$  are insignificantly correlated with regional market returns (at any frequency), or regional variants of the size ( $SMB$ ), value-growth ( $HML$ ) and momentum ( $WML$ ) factors at the monthly level (available from Kenneth French's website).

#### **4.6 FURTHER EVIDENCE OF MISVALUATION VS. INFORMATION DIFFUSION**

Thus far I have shown that twin return differentials contain a strong common factor at the country level. This common factor, denoted the local habitat factor, is hypothesized to reflect

misvaluation. An alternative explanation is that this commonality in twin return differentials reflects differences in information diffusion, either non-trading mechanically induces a spurious correlation between return differentials among ETF-pairs from the same country-pairs with similar non-trading probabilities, or local investors are better informed about the local economy relative to foreign investors and therefore impound fundamental news faster into the prices of locally traded ETFs.

In the previous section I already provided some evidence against the alternative explanation of information diffusion by showing that the correlation between ETF misvaluation and local stock market returns declines rapidly with the return horizon, consistent with misvaluation. In this section I provide further evidence based on three tests.

First, I analyze how the degree of commonality in return differentials is related to the liquidity characteristics and arbitrage costs of ETF-pairs. According to the alternative explanation of information diffusion, commonality in return differentials should be stronger for less liquid ETFs that have a higher degree of non-trading (or stale pricing) probability. The prediction is less clear if commonality in return differentials reflects local-trading induced misvaluation. Previous studies suggest that commonality in return differentials should be stronger during periods of market-wide illiquidity and limits to arbitrage (Kumar and Lee, 2006; Kumar and Spalt, 2013). Alternatively, in the cross-section of securities, demand is likely to go first to the most liquid securities, implying that commonality in return differentials could be stronger for more liquid securities. Previous findings also suggest that liquid securities are more attractive to short-term noise traders, which can make these assets more exposed to non-fundamental demand shocks (as argued in Chapter 2).

Second, I provide an experiment that attempts to remove any local information advantage that local investors may have by focusing on ETFs traded in countries other than the trading location of their underlying securities. In this case commonality in return differentials is highly unlikely to reflect differences in information diffusion.

Third, I test an important implication of the local-trading induced misvaluation hypothesis, namely that premiums of the local ETF (relative to NAV) should predict future returns of the local ETF negatively, as opposed to premiums of the foreign ETF (relative to NAV) predicting future returns of the local ETF positively as it catches up to reflect the fundamental information reflected in foreign ETF prices.

#### 4.6.1 Liquidity and commonality in misvaluation

Several liquidity measures have been proposed in the literature, but many of them are not comparable across countries due to differences in trading mechanisms and conventions, especially measures related to trading volume (e.g. Amihud's price impact and turnover). Instead, I use the simplest and perhaps the most widely used measure of liquidity for normalized transactions, which is the proportional quoted spread, defined as the difference between the closing ask and bid prices divided by the mid-quote prices, averaged over a monthly period.

My second measure is an implicit measure of liquidity, namely arbitrage costs. Idiosyncratic risk is one of the most important measures of arbitrage costs. Pontiff (2006) demonstrates theoretically that a rational investor's demand for a mispriced asset increases with the magnitude of misvaluation, but decreases with the asset's idiosyncratic risk. His argument relies on the idea that the hedge position intended to capture the misvaluation must, by design, be uncorrelated with the market and any other hedge positions. An arbitrageur will therefore be exposed to idiosyncratic risk whenever she has to delay liquidating the position<sup>53</sup>. Empirical evidence by Gagnon and Karolyi (2010) in the context of cross-listed stocks also demonstrates the importance of idiosyncratic risk for misvaluation. To measure idiosyncratic risk, I use the standard deviation of ETF-NAV return (Std. Dev. of the change in premium in Eq. (3)) measured over a monthly period.

As with the liquidity measures, I estimate the degree of commonality in twin return differentials using daily data over a calendar month. For the estimation I require valid (non-missing) data for a minimum of 75 percent of trading days. Commonality in twin return differentials is measured either from a regression of twin return differentials on the composite local habitat factor (Eq. (6)), or from a regression of twin return differentials on the local stock market returns of countries  $L$  and  $F$ . In either case I also control for the lagged level of ETF premiums.

---

<sup>53</sup> In the context of ETFs, limited delays can occur because share creations only take place at end-of-day NAVs, while the underlying portfolio may have to be accumulated over an extended period of time for several reasons: 1) the APs need to hedge their exposure to the underlying securities when an ETF is sold in the secondary market until enough demand is available to meet the minimum creation unit size, 2) the additional costs and risks faced by cross-border arbitrage trades may deter arbitrage activity until mispricing is large enough, 3) arbitrageurs may wish to avoid price impact costs by splitting up the purchases over an extended period of time. Moreover, search and delay costs are more likely to arise when the position to be liquidated is large.

In this context, how should one measure the *degree* of commonality? The existing literature mainly uses the regression  $R^2$  (e.g. Hameed, Kang and Viswanathan, 2010; Karolyi, Lee and Van Dijk, 2012), although the beta coefficient is also used (Kamara, Lou and Sadka, 2008). Unlike the  $R^2$ -measure, the beta coefficient is sensitive to scaling effects that arise from differences in factors and their volatilities (i.e. the beta denominator) across ETFs in different countries (Brockman *et al.*, 2009). For this reason I focus primarily on the  $R^2$ -measure, although for robustness I report results for both. In order to remove the influence of the lagged twin price differential on model  $R^2$ , I perform a decomposition as follows:

$$R_i^2 = \frac{COV(R_{i-p,t}, \hat{\alpha}_{1,i} P_{i-p,t-1})}{VAR(R_{i-p,t})} + \frac{COV(R_{i-p,t}, \beta_{L-F,i} R_{L-F,t})}{VAR(R_{i-p,t})} \quad (29)$$

I use the *second* normalized covariance term, denoted  $R_{L-F,i}^2$ , to measure the degree of commonality in twin return differentials for a given ETF-pair and month.  $R_{L-F,i}^2$  can be interpreted as the fraction of the model  $R^2$  attributable to the local habitat factor (Graham, Li and Qiu, 2013). I perform a similar decomposition for the second regression that uses local stock market returns instead of the composite habitat factor:

$$R_i^2 = \frac{COV(R_{i-p,t}, \lambda_i P_{i,t}^{E-peer})}{VAR(R_{i-p,t})} + \frac{COV(R_{i-p,t}, \beta_{L,i} R_{MKT,t}^L)}{VAR(R_{i-p,t})} + \frac{COV(R_{i-p,t}, \beta_{F,i} R_{MKT,t}^F)}{VAR(R_{i-p,t})} \quad (30)$$

I use the sum of the second and third normalized covariance terms, denoted  $R_{MKT,i}^2$ , as my second measure for commonality in twin return differentials.

Table 4-8 provides estimates for the average degree of commonality in twin return differentials ( $R_{L-F}^2$  or  $\beta_{L-F}$ ) using a two-way sort on the prior month liquidity of ETF  $i$  (long-leg of misvaluation), and the liquidity of its foreign twin  $p_i$  (short-leg). This allows us to compare the degree of commonality in twin return differentials among ETFs, holding constant the fundamentals (i.e. identical underlying benchmark indices). There are three liquidity terciles each for ETF  $i$ , and  $p_i$ , which yields nine liquidity groupings in total. To be included in the sample, I require a minimum of 9 foreign peers per ETF, or at least one per liquidity group. This is not a serious restriction because the average (median) ETF has 94 (18) foreign peers.

[Table 4-8]

The results in Table 4-8 suggest that if we hold the liquidity of ETF  $i$  constant (long-leg), commonality in twin return differentials increases with the liquidity of the foreign peer  $p_i$  (short-leg). The results hold for both measures of liquidity (proportional quoted spreads and idiosyncratic risk) as well as for both measures of commonality ( $R_{L-F}^2$  and  $\beta_{L-F}$ ). Holding ETF  $i$ 's idiosyncratic risk constant at high:  $R_{L-F}^2$  increases from 28.6 to 40.2 to 62.9 percent as the idiosyncratic risk of the short-leg decreases from high to medium to low. These increases are highly significant on any conventional level. Similar, albeit less pronounced and significant increases are observed when the liquidity of the long-leg is fixed at medium or high. The increase in commonality for more liquid funds in the short-leg is consistent with the local-trading induced misvaluation story where demand shocks materialize first in the most liquid ETFs, but the results are opposite to the predictions of the information diffusion story.

Overall, commonality in twin return differentials is most pronounced when the liquidity of the short-leg is high, while there is little difference in the degree of commonality when the liquidity of the long-leg increases. This asymmetry – liquidity matters only for the ETF in the short-leg – may be related to country-specific characteristics in the sample. Since every ETF-pair is unique, the ordering by country-pairs is arbitrary. In my sample U.K. and Switzerland combined account for more than half of the sample of ETFs in the long-leg, while Italy and Netherlands account for roughly half of the sample of ETFs in the short-leg<sup>54</sup>.

#### 4.6.2 No information advantage

In this section I provide an experiment where local investors have no informational advantage, in which case commonality in twin return differentials is highly unlikely to reflect differences in information diffusion. Specifically, I estimate regression of twin return differentials on the local habitat factors for two sub-samples. The first sample contains ETF-pairs where the trading location of one of the ETFs matches the trading location of the underlying securities (e.g. pair of ETFs tracking FTSE 100, where one of the pairs is traded in the U.K., the other is traded in France). The second sample contains ETF-pairs where the trading location of both ETFs is

---

<sup>54</sup> I start the matching from U.K., followed by SWI, FRA, GER, ITA and NET. In my case U.K. is the first country in the long-leg and therefore never appears in the short-leg (as every ETF-pair consists of an ETF from different countries). For example, the most popular benchmark index is Eurostoxx 50 with 78 ETFs tracking this index. When forming pairs of ETFs, I start with U.K. listed ETF. I match this ETF, one-by-one, with every identical ETF traded in Switzerland, then I move on to the next country (France), followed by the next (Germany) and so on. The sequence of countries is fixed (UK, SWI, FRA, GER, ITA, NET), which is why U.K. appears most frequently in the long-leg, while Italy and Netherlands appear most often in the short-leg.

separated from the trading location of the underlying securities. For instance, pairs of ETFs tracking CAC 40 where both pairs are traded in countries other than France. If commonality in twin return differentials is driven by information diffusion, then we should expect the results to be stronger for the first sample relative to the second. The regressions are otherwise identical to those estimated before using regression (7) with the composite local habitat factor, or regression (10) with separate habitat factors for the local markets of countries  $L$  and  $F$ .

[Table 4-9]

The results in Table 4-9 reveal no meaningful differences in the degree of commonality in twin return differentials between the two samples, whether at the daily or monthly horizons, and regardless of the regression specification used. The absence of results supports my conjecture that commonality in twin return differentials is in fact driven by misvaluation, and not information diffusion.

Although the sample size used in this experiment is only 11 percent of the full sample, the number of daily ETF-pair observations is still considerable at more than 200,000. Nevertheless, I also conduct a similar test of no information advantage by using sub-samples of ETFs tracking either single-country indices vs. Pan-European indices. Local investors are unlikely to have an informational advantage over regional equities. These results (unreported for conciseness) also do not reveal any economically significant differences between the two samples.

### **4.6.3 Return predictability and the reversal of misvaluation**

An important testable implication of local-trading induced misvaluation is its subsequent reversal as arbitrage forces restore pricing efficiency. Let us assume that the ETF  $i$  is hit by a positive local demand shock that pushes its price above the fundamental NAV value ( $PREM_{i,t} > 0$ ), while its twin security ETF  $p_i$  is not affected (because it is owned by foreign investors) and remains correctly priced ( $PREM_{p_i,t} = 0$ ). If  $PREM_{i,t}$  truly reflects misvaluation of the ETF, then we should observe a reversal in the future returns of ETF  $i$  ( $R_{i,t+\tau} < 0$ ) with no impact on future returns of ETF  $p_i$  ( $R_{p_i,t+\tau} = 0$ ) in the absence of news to fundamentals.

The alternative hypothesis is that the initial demand shock was driven by positive fundamental news. In this case ETF  $i$  is correctly valued and its future returns are unaffected in the absence of other fundamental news ( $R_{i,t+\tau} = 0$ ). In contrast, future returns for the twin

security  $p_i$  will be positive as it catches up to reflect the fundamental news already incorporated in the price of ETF  $i$  ( $R_{p_i,t+\tau} > 0$ ). To test these implications, I estimate the following regression:

$$R_{i,t+\tau} = a_i + b_i PREM_{i,t} + c_i PREM_{p_i,t} + e_{i,t+\tau} \quad (31)$$

$$R_{p_i,t+\tau} = a_{p_i} + b_{p_i} PREM_{i,t} + c_{p_i} PREM_{p_i,t} + e_{p_i,t+\tau} \quad (32)$$

where  $R_{i,t+\tau}$  = Return measured over the interval  $t+\tau$  for ETF  $i$ ,  $i \in L$   
 $R_{p_i,t+\tau}$  = Return measured over the interval  $t+\tau$  for ETF  $p_i$ ,  $p_i \in F$   
 $PREM_{i,t}$  = ETF premium relative to NAV on day  $t$

The LHS and RHS variables are standardized to obtain coefficient estimates that are comparable across various return intervals ( $\tau$ ). If  $PREM_{i,t}$  reflects misvaluation for ETF  $i$ , and  $PREM_{p_i,t}$  reflects misvaluation for ETF  $p_i$ , then a future reversal in misvaluation implies that  $b_i < 0$  and  $c_{p_i} < 0$ . In contrast, the information diffusion story predicts that  $PREM_{i,t}$  reflects fundamental news already incorporated into the price of ETF  $i$ , implying that future returns of  $p_i$  will catch up with a lag ( $b_{p_i} > 0$ ). Similarly, when  $PREM_{p_i,t}$  reflects fundamental news incorporated in  $p_i$ , then future returns of ETF  $i$  will catch up with a lag ( $c_i > 0$ ).

In order to account for the dependence in residuals across funds (because we have not accounted for common fundamental risk) and possibly also over time, I estimate regressions (15-16) using pooled OLS with Driscoll and Kraay (1998) standard errors that are robust to general forms of cross-sectional and time-series dependence. The maximum lag order of autocorrelation is set to the number of overlapping days per return horizon.

[Table 4-10]

The results in Table 4-10 show that an ETFs own premium predicts future returns negatively ( $b_i < 0$  and  $c_{p_i} < 0$ ) consistent with the reversal of local demand shocks. As the data is standardized, the magnitude of the coefficient estimates can be interpreted as percentages of Std. Dev. of the dependent variable. Thus, a one Std. Dev. increase in an ETFs own premiums is associated with a 14.9 percent and 13.4 percent decline in returns for ETF  $i$  and  $p_i$  respectively over a daily period. The magnitudes decline by roughly half over a weekly horizon, and by two thirds over a monthly horizon. Nevertheless, the coefficients are significant at least over a return horizon of six months.

In contrast to local trading induced mispricing, the cross-terms are positive implying that current misvaluation of ETF  $i$  predicts future returns of ETF  $p_i$  positively ( $b_{p_i} > 0$ ), and current misvaluation of ETF  $p_i$  predicts future returns of  $i$  positively ( $c_i > 0$ ). Economically these results are much weaker as the coefficient estimates for the cross-terms are less than half as strong as those for the own-terms, and the cross-terms are generally significant only for return horizons up to a week. Another interpretation is that the return predictability partially reflects arbitrage activity. For instance, when ETF  $i$  is trading at a premium relative to NAV and relative to  $p_i$ , then arbitrageurs would buy the underpriced asset (ETF  $p_i$ ) and sell the overpriced asset (ETF  $i$ ), which would induce a reversal in the future returns of ETF  $i$  and an appreciation of the future returns of ETF  $p_i$ . With this interpretation it may be more useful to investigate the difference in coefficient estimates for an ETFs own premium and the premium of its peer  $p_i$ . Since the cross-terms are always weaker by half, the results continue to support the reversal of local-trading induced mispricing.

#### **4.7 SUMMARY AND CONCLUSIONS**

In this study I investigate whether differences in local clienteles between identical ETF-pairs from different countries can be used to identify commonality in misvaluation among locally traded securities. My sample contains 416 plain-vanilla Exchange-Traded Funds from six European countries tracking European equity indices; the total number of ETF-pairs is 2727. The theoretical justification is based on the “preferred habitat” theory of Barberis, Shleifer and Wurgler (2005) where some securities (locally traded ETFs) are predominantly owned and traded by a specific subset of investors (local investors).

This paper expands on existing research in several ways. First, local clienteles are clearly identified because the primary motivation for listing new (or cross-list existing) funds is to support local clientele demand, according to ETF providers. Second, because trading hours are almost perfectly aligned in Europe and because ETF misvaluation can be calculated relative to i) own (cross-listed) shares traded abroad, ii) foreign peers tracking the same underlying benchmark passively or iii) underlying portfolio NAV, I can control for not only non-trading, but also stale pricing.

My findings indicate that daily twin return differentials (Local-Foreign) comove strongly with a local habitat factor constructed from the equally-weighted twin return differential of other

ETF-pairs in the same country-pairs  $L$  and  $F$ . The degree of commonality in return differentials is also considerable when compared against the variability in raw returns (15.34 percent), but declines rapidly with the return horizon (only 3.75 percent in monthly horizons), consistent with misvaluation.

I conduct several tests to rule out the alternative explanation that commonality in return differentials reflects (fundamental) differences in information diffusion between countries. First, I compare the degree of commonality in return differentials for two sub-samples: a) a sample where the trading location of ETF-pairs is different from to the trading location of their underlying securities with b) a sample where the trading locations of the ETF-pairs and their underlying securities match. The results do not materially differ between the two samples. Third, I compare the degree of commonality in twin return differentials among ETFs with identical fundamentals, but different levels of liquidity (and non-trading probability). In contrast to the information diffusion argument which predicts that commonality in return differentials should increase with illiquidity and non-trading probability, my findings indicate the opposite.

Consistent with the conjecture that both locally traded ETFs and stocks are exposed to local-trading induced misvaluation, I find that twin return differentials are positively correlated with stock market returns in the local market, and negatively with the returns in the foreign market. This relationship is strongest at the daily horizon, weaker but still significant at weekly horizon, and disappears entirely at the monthly horizon.

Overall, my findings indicate that local clientele and the trading by local investors can generate commonality in misvaluation.

**Table 4- 1: Selecting ETF twins**

This table provides several examples to illustrate how ETF twins are selected. Panel A shows a randomly selected ETF (ETF  $i$ ), while Panel B shows the potential twin securities for ETF  $i$ . ID is a unique identifier for a given ETF and trading location. The variable ticker includes the ticker of the security and the exchange country (as reported by Bloomberg). Listing type is either primary (where original IPO took place), or cross-listing. FX is the currency in which the ETF shares are traded (GBX = penny sterling, or GBP/100). There are three different methods to select twins. First, we can use an ETFs own shares traded in a foreign market. Second, we can use the shares of a different ETF with an identical benchmark index traded in a foreign country or third, traded in the same country.

Name	ID	Ticker	Exchange	Listing type	FX	ETF-pair ID (by twin type)		
						1: Foreign (own)	2: Foreign (different)	3: Domestic
<b>PANEL A: ETF <math>i</math></b>								
iShares Euro Stoxx 50	1	EUEA NA	London	Cross	GBX			
<b>PANEL B: twin ETF <math>p_i</math></b>								
iShares Euro Stoxx 50	2	EUE IM	Borsa Italian	Cross	EUR	Pair: 1, 2	Pair: 1, 2	
	3	EUE LN	Euronext Amst.	Primary	EUR	Pair: 1, 3	Pair: 1, 3	
	4	EUN2 GR	Xetra ETF	Cross	EUR	Pair: 1, 4	Pair: 1, 4	
	5	EUNE SW	SIX	Cross	CHF	Pair: 1, 5	Pair: 1, 5	
LYXOR ETF Euro Stoxx 50	6	LYSX GR	Xetra ETF	Cross	EUR		Pair: 1, 6	
	7	MSE FP	Euronext Paris	Primary	EUR		Pair: 1, 7	
AMUNDI ETF Euro Stoxx 50 C	8	C50 FP	Euronext Paris	Primary	EUR		Pair: 1, 8	
	9	C50 IM	Borsa Italian	Cross	EUR		Pair: 1, 9	
HSBC Euro Stoxx 50 ETF	10	H50E LN	London	Cross	GBX			Pair: 1, 10
	11	50E FP	Euronext Paris	Primary	EUR		Pair: 1, 11	

**Table 4- 2: Snapshot of ETF sample 11/2013**

This table provides a snapshot of the Assets Under Management (AUM) at the beginning (01/2006), middle (01/2010) and end of the sample (11/2013). The number of ETFs is reported at the end of the sample. The number of ETFs is reported both for primary listings (i.e. an ETF's first, or original listing), as well as the total number of listings (including both original and cross-listings). AUM is measured as the number of shares outstanding multiplied by the ETF share price in the market of the original IPO.

Sample	AUM (in \$million)			Number of ETF listings	
	01/2006	01/2010	11/2013	Primary	Total
All	22,563	86,072	140,800	186	416
<b>By replication method</b>					
Synthetic	9,373	32,425	42,474	82	185
Sample		8,474	16,479	19	55
Full	13,190	45,174	81,848	85	176
<b>By Trading Location</b>					
UK				28	62
SWI				30	87
FRA				46	52
GER				79	126
ITA				1	71
NET				2	18
<b>By Investment Location</b>					
UK	65	8,661	16,099	25	52
SWI	2,521	6,431	9,622	15	22
FRA	5,000	6,224	6,327	7	12
GER	3,352	9,658	31,766	11	19
ITA	264	570	2,739	5	14
NET		607	569	4	5
Country	11,202	32,152	67,122	67	124
Regional	11,360	53,921	73,678	119	292
Europe Large-Cap	10,956	49,696	66,336	85	212
Europe Mid-Cap	175	666	1,594	9	19
Europe Small-Cap	73	1,766	3,246	9	23
Other	156	1,793	2,502	16	38

**Table 4- 3: Summary statistics on ETF misvaluation**

Panel A provides summary statistics for the level of (change in) premiums, or the log-price (return) differential between an ETF and its underlying portfolio NAV. Panel B, C and D provide summary statistics for the log-price (return) differential between pairs of twin ETFs. In panel B, the twin of an ETF refers to the same ETF share traded in a foreign market. In Panel C and D the twin refers to a different ETF with an identical benchmark index as ETF  $i$  and which is traded in the foreign (domestic) market relative to  $i$ . In this case the price levels are not directly comparable, which is why I scale the ETF prices by their corresponding NAV prices before taking the log-price difference (this corresponds to taking the difference in ETF premiums relative to NAV). All variables are measured in percentage. Summary statistics are calculated over the 01/2006-11/2013 period (weekly data) separately for four different samples: a) all funds, b) funds that use synthetic replication, c) sample replication or d) full replication.

Variable	Sample	Mean	Median	Std. Dev.	1 %	99 %
<b>Panel A: Twin = underlying portfolio NAV</b>						
Premium: $PREM_{i,t} = \ln(P_{i,t}/NAV_{i,t})$	All	-0.004	-0.005	0.275	-0.803	0.792
$\Delta$ Premium: $\Delta PREM_{i,t}$	All	0.006	0.002	0.644	-1.964	2.040
<b>Panel B: Twin = same ETF share traded abroad</b>						
Nr. of ETF pairs	All	3.670	3.000	2.746	1.000	10.000
Twin price differential:	All	0.027	0.014	0.295	-0.771	0.912
$P_{i-p1,t} = \ln(P_{i,t}/P_{p1,t})$	Synthetic replication	0.025	0.015	0.181	-0.438	0.565
	Sample replication	0.019	0.009	0.171	-0.451	0.580
	Full replication	0.031	0.015	0.197	-0.405	0.670
Twin return differential:	All	0.015	0.002	0.964	-3.077	3.112
	Synthetic replication	0.034	0.011	1.483	-4.375	4.311
	Sample replication	0.132	0.031	1.606	-3.906	5.239
	Full replication	0.067	0.017	1.374	-4.065	4.315
<b>Panel C: Twin = different ETF (foreign traded) with an identical benchmark index</b>						
Nr. of ETF pairs	All	94.035	18.000	239.161	4.000	1065.000
Twin price differential:	All	-0.005	-0.006	0.243	-0.604	0.602
$P_{i-p2,t} = PREM_{i,t} - PREM_{p2,t}$	Synthetic replication	-0.005	-0.007	0.226	-0.551	0.606
	Sample replication	0.000	0.000	0.187	-0.540	0.527
	Full replication	0.003	-0.002	0.265	-0.596	0.617
Twin return differential:	All	0.018	0.003	0.962	-3.033	3.095
	Synthetic replication	0.015	0.004	1.010	-3.358	3.299
	Sample replication	0.023	0.001	0.913	-2.851	2.971
	Full replication	0.017	0.003	0.968	-2.943	3.024
<b>Panel D: Twin = different ETF (domestic traded) with an identical benchmark index</b>						
Nr. of ETF pairs	All	27.478	6.000	63.675	1.000	261.000
Twin price differential:	All	0.017	0.007	0.202	-0.517	0.537
$P_{i-d,t} = PREM_{i,t} - PREM_{d,t}$	All	0.017	0.007	0.202	-0.517	0.537
Twin return differential:	All	-0.009	-0.001	0.694	-2.235	2.219
$R_{i-d,t} = R_{i,t} - R_{d,t}$	All	-0.009	-0.001	0.694	-2.235	2.219

**Table 4- 4: Trading location specific misvaluation**

This table reports the results from the following regression:

$$R_{i-p,t} = \alpha_{0,i} + \alpha_{1,i} P_{i-p,t-1} + \beta_{L-F,i} R_{L-F,t} + \varepsilon_{i,t}$$

where  $P_{i-p,t}$  ( $R_{i-p,t}$ ) is the log price (return) twin return differential based on the difference between 1) an ETF and its (own) shares traded in a foreign market or 2) ETF  $i$  and a different ETF with an identical benchmark index traded in a foreign market. The local-minus-foreign habitat factor  $R_{L-F,t}$  is calculated as the equally-weighted twin return differential of other ETF-pairs traded in the same country-pairs as ETF  $i$  and its foreign twin  $p_i$ . Coefficients are shown for the average, the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentile of the distribution. T-statistics for the mean are adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). T-statistics at the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> refer to fund specific t-statistics and are based on white standard errors. Avg. R<sup>2</sup> is the average R-squared across all funds, Nobs is the number of observations. Impact (basis points) is the impact of a 1 Std. Dev. shock increase in the factor on the dependent variable, or  $\beta_{L-F} * Std(R_{L-F,t})$ , while Impact (% of SD) is the proportional impact, or  $\beta_{L-F} * STD(R_{L-F,t})/STD(R_{i-p,t})$

Sample	$\alpha_1$	[t-stat]	$\beta_{L-F}$	[t-stat]	Impact (bp)	Impact (% SD)	Avg. R <sup>2</sup>	N obs
<b>PANEL A: Twin = own ETF shares traded in a foreign market</b>								
<b>Daily</b>								
Average	-0.614	[73.28]	0.596	[76.33]	0.206	0.411	0.561	491,319
P05	-0.924	[16.60]	0.209	[2.22]				
Median	-0.634	[9.28]	0.589	[9.58]				
P95	-0.237	[4.70]	1.071	[32.19]				
<b>Monthly</b>								
Average	-0.522	[18.82]	0.652	[19.54]	0.272	0.437	0.470	22,351
P05	-1.145	[5.83]	-0.040	[0.10]				
Median	-0.506	[2.28]	0.665	[3.25]				
P95	0.006	[0.02]	1.269	[9.35]				
<b>PANEL B: Twin = a different ETF with an identical benchmark index</b>								
<b>Daily</b>								
Average	-0.454	[64.84]	0.727	[103.05]	0.250	0.543	0.633	2,481,573
P05	-0.818	[15.29]	0.247	[3.61]				
Median	-0.430	[8.20]	0.707	[15.33]				
P95	-0.176	[4.84]	1.178	[39.43]				
<b>Monthly</b>								
Average	-0.537	[18.25]	0.691	[23.07]	0.277	0.494	0.569	112,982
P05	-1.131	[5.99]	0.048	[0.18]				
Median	-0.483	[2.63]	0.727	[4.15]				
P95	-0.109	[0.58]	1.252	[9.88]				

**Table 4- 5: Separating the local and foreign habitat factors**

This table reports the results from the following regression:

$$R_{i-p,t} = \alpha_{0,i} + \alpha_{1,i}P_{i-p,t-1} + \beta_{L,i}R_{Local,t} + \beta_{F,i}R_{Foreign,t} + \beta_{R,i}R_{REST,t} + \varepsilon_{i-p,t}$$

where  $P_{i-p,t}$  ( $R_{i-p,t}$ ) is the log price (return) twin return differential based on the difference between 1) ETF  $i$  and its (own) shares traded in a foreign market or 2) ETF  $i$  and a different ETF with an identical benchmark index traded in a foreign market. The local habitat factor  $R_{Local,t}$  is calculated as the equally-weighted ETF-NAV return differential of other ETFs in the same country as ETF  $i$  (long-leg), while the foreign habitat factor  $R_{Foreign,t}$  is calculated as the equally-weighted ETF-NAV return of other ETFs traded in the same country as the ETFs peer (short-leg).  $R_{REST,t}$  is the equally-weighted ETF-NAV return difference of ETFs traded outside the two markets. Coefficients are shown for the average, the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentile of the distribution. T-statistics for the mean are adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). T-statistics at the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> refer to fund specific t-statistics and are based on white standard errors. A.  $R^2$  is the average R-squared across all funds,  $N$  is the number of observations. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % levels respectively.

Sample	$\beta_L$	[t-stat]	$\beta_F$	[t-stat]	$\beta_R$	[t-stat]	A. $R^2$	$N$
<b>PANEL A: Twin = own ETF shares traded in a foreign market</b>								
<b>Daily</b>								
Average	0.540	[35.50]	-0.569	[53.11]	0.036	[2.37]	0.566	483,897
P05	0.106	[0.60]	-1.025	[26.26]	-0.284	[2.29]		
Median	0.515	[3.35]	-0.566	[11.85]	0.005	[0.97]		
P95	1.058	[5.91]	-0.171	[7.20]	0.458	[0.07]		
<b>Monthly</b>								
Average	0.443	[7.89]	-0.582	[11.23]	-0.012	[0.17]	0.488	22,351
P05	-0.387	[0.69]	-1.235	[8.00]	-0.971	[2.21]		
Median	0.471	[0.44]	-0.613	[4.06]	-0.072	[1.04]		
P95	1.135	[1.50]	0.122	[2.52]	1.164	[0.26]		
<b>PANEL B: Twin = a different ETF with an identical benchmark index</b>								
<b>Daily</b>								
Average	0.680	[40.39]	-0.707	[61.33]	0.016	[1.32]	0.659	2,089,110
P05	0.164	[1.45]	-1.221	[37.93]	-0.197	[2.33]		
Median	0.720	[5.37]	-0.733	[20.09]	0.004	[1.02]		
P95	1.165	[9.38]	-0.182	[10.50]	0.305	[0.10]		
<b>Monthly</b>								
Average	0.545	[8.82]	-0.673	[14.78]	-0.006	[0.14]	0.603	95,038
P05	-0.211	[0.58]	-1.235	[6.85]	-0.428	[1.88]		
Median	0.578	[0.94]	-0.712	[4.83]	-0.006	[0.80]		
P95	1.262	[2.16]	0.039	[3.32]	0.398	[0.03]		

**Table 4- 6: Placebo test for local misvaluation in ETF returns**

This table reports the results from the following regression :

$$R_{i-d,t} = \alpha_{0,i} + \alpha_{1,i}P_{i-d,t-1} + \beta_{L,i}R_{Local,t} + \beta_{R,i}R_{REST,t} + \varepsilon_{i-d,t}$$

where  $P_{i-p,t}$  ( $R_{i-p,t}$ ) is the log price (return) twin return differential based on the difference pairs of ETFs with identical benchmark index traded in the same country. The local habitat factor  $R_{Local,t}$  is calculated as the equally-weighted ETF-NAV return differential of other ETFs in the same country as ETF  $i$  (long-leg).  $R_{Rest,t}^{E-N}$  is the equally-weighted ETF-NAV return difference of ETFs traded outside the local market. Coefficients are shown for the average, the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentile of the distribution. T-statistics for the mean are adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). T-statistics at the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentile refer to fund specific t-statistics and are based on white standard errors. A.  $R^2$  is the average adjusted R-squared,  $N$  is the number of observations. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % levels respectively.

Sample	$\alpha_1$	[t-stat]	$\beta_L$	[t-stat]	$\beta_R$	[t-stat]	A. $R^2$	$N$
<b>Daily</b>								
Average	-0.011	[2.38]	0.004	[0.62]	-0.011	[2.38]	0.381	485,379
P05	-0.327	[5.78]	-0.092	[2.37]	-0.327	[5.78]		
Median	-0.002	[0.14]	-0.001	[-0.04]	-0.002	[0.14]		
P95	0.221	[4.10]	0.142	[2.40]	0.221	[4.10]		
<b>Monthly</b>								
Average	-0.005	[0.25]	0.007	[0.37]	-0.005	[0.25]	0.453	22,622
P05	-0.298	[1.68]	-0.251	[1.67]	-0.298	[1.68]		
Median	0.007	[0.09]	0.003	[0.03]	0.007	[0.09]		
P95	0.237	[1.71]	0.274	[1.60]	0.237	[1.71]		

**Table 4- 7: Twin return differentials and local stock market returns**

This table reports the results from regressions of twin return differentials (of ETF  $i$  relative to its own shares traded abroad) on country-specific risk factors and the lagged log price differential (omitted for conciseness). Local and foreign stock market returns are value-weighted returns across all stocks traded in the local market (of ETF  $i$ ) and the foreign market (of ETF  $i$ 's foreign twin). The country-specific size ( $SMB$ ) and value ( $HML$ ) factors are calculated similar to Fama and French (2012) with the exception that the cut-offs are country specific. The  $P\&S\ LIQ$  factor corresponds to the innovation in country specific liquidity, calculated as in Pastor and Stambaugh (2003). The average coefficients are reported. T-statistics for the mean are adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). Avg.  $R^2$  is the average R-squared,  $N$  is the number of observations. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % levels respectively.

Variables	Daily		Weekly		Monthly		
	(1)	(2)	(1)	(2)	(1)	(2)	(3)
<b>PANEL A: Unit of observation = ETF-pair</b>							
<i>VW-MKT</i> (Local)	0.064*** (9.76)	0.068*** (9.53)	0.027*** (4.21)	0.038*** (5.33)	-0.009 (-1.02)	-0.007 (-0.60)	-0.011 (-0.91)
<i>VW-MKT</i> (Foreign)	-0.056*** (10.48)	-0.056*** (6.57)	-0.019*** (3.69)	-0.029*** (4.37)	0.001 (0.19)	0.002 (0.17)	0.004 (0.42)
<i>SMB</i> (Local)		-0.001 (0.23)		0.013* (1.89)		-0.020* (1.87)	-0.021* (1.95)
<i>SMB</i> (foreign)		0.008 (0.68)		-0.006 (0.82)		0.000 (0.02)	-0.002 (0.17)
<i>HML</i> (Local)		-0.007 (1.2)		-0.017*** (2.79)		-0.014 (1.31)	-0.013 (1.27)
<i>HML</i> (Foreign)		0.009 (1.41)		0.013*** (2.10)		0.008 (0.99)	0.009 (1.08)
<i>P&amp;S LIQ</i> (Local)							-2.540 (0.67)
<i>P&amp;S LIQ</i> (Foreign)							-3.408 (1.78)
Avg. $R^2$	0.43	0.43	0.38	0.40	0.30	0.38	0.41
$N$	491,887	491,845	100,246	100,246	22,764	22,764	22,764
<b>PANEL B: Unit of observation = Country-pair</b>							
<i>VW-MKT</i> (Local)	0.058*** (13.21)	0.060*** (11.29)	0.020*** (3.91)	0.021*** (3.86)	0.005 (1.08)	0.001 (0.15)	0.002 (0.32)
<i>VW-MKT</i> (Foreign)	-0.048*** (12.11)	-0.050*** (10.36)	-0.017*** (3.69)	-0.016*** (3.40)	-0.012*** (2.39)	-0.005 (0.92)	-0.005 (0.91)
<i>SMB</i> (Local)		0.001 (0.13)		-0.002 (0.48)		-0.009* (1.72)	-0.009* (1.72)
<i>SMB</i> (foreign)		-0.001 (0.23)		0.004 (0.92)		0.005 (1.21)	0.005 (1.09)
<i>HML</i> (Local)		-0.003 (0.62)		-0.005 (1.16)		0.005 (0.89)	0.005 (0.94)
<i>HML</i> (Foreign)		0.005 (1.37)		0.002 (0.73)		-0.005 (1.26)	-0.005 (1.12)
<i>P&amp;S LIQ</i> (Local)							0.299 (0.15)
<i>P&amp;S LIQ</i> (Foreign)							-1.714 (1.37)
Avg. $R^2$	0.43	0.43	0.38	0.39	0.26	0.30	0.32
$N$	26,668	26,664	5,667	5,667	1,304	1,304	1,304

**Table 4- 8: Degree of commonality in twin return differentials and fund-pair liquidity**

The first set of columns (1 and 2) reports comovements from a regression of twin return differentials (between ETF  $i$  and its own shares traded abroad) on the local-minus-foreign habitat factor (see Table 4). The second set of columns (3 and 4) reports comovements from a regression of twin return differentials on the local and foreign stock market returns (see Table 7). The degree of commonality in twin return differentials is measured either as the fraction of R-squared attributable to the local-minus-foreign factor or ii) the beta (or regression) coefficient on the local-minus-foreign factor, using daily data with at least 15 observations per month. Results are reported for sub-samples of ETF pairs sorted by their liquidity (quoted spread in Panel A; idiosyncratic risk in Panel B).

		Measures of comovement							
		(1): $R_{L-F}^2$		(2): $\beta_{L-F}$		(3): $R_{MKT}^2$		(4): $\beta_{MKT}$	
		$\Delta$ Mean		$\Delta$ Mean		$\Delta$ Mean		$\Delta$ Mean	
Long-leg	Short-leg	Mean	[t-stat]	Mean	(t-stat)	Mean	(t-stat)	Mean	(t-stat)
<b>PANEL A: Sorting variable = Quoted Spread</b>									
High	High	47.554		0.692		11.580		0.157	
High	Med	50.706	[7.41]	0.720	[4.84]	11.962	[2.68]	0.192	[7.63]
High	Low	54.173	[8.53]	0.777	[10.20]	12.331	[2.62]	0.239	[9.66]
Med	High	43.492		0.672		10.721		0.126	
Med	Med	44.983	[4.40]	0.678	[1.15]	11.263	[4.76]	0.155	[8.03]
Med	Low	46.963	[5.80]	0.724	[9.37]	11.680	[3.65]	0.202	[11.75]
Low	High	43.969		0.712		10.694		0.147	
Low	Med	45.849	[4.69]	0.719	[1.22]	11.242	[4.15]	0.184	[8.06]
Low	Low	49.316	[8.17]	0.777	[9.51]	11.473	[1.66]	0.218	[6.53]
<b>PANEL B: Sorting variable = Idiosyncratic Risk</b>									
High	High	28.589		0.476		9.273		0.033	
High	Med	40.216	[30.22]	0.606	[21.26]	11.325	[14.90]	0.140	[29.71]
High	Low	62.952	[61.43]	0.853	[47.49]	13.818	[18.21]	0.326	[39.57]
Med	High	33.539		0.615		9.266		0.031	
Med	Med	38.320	[15.92]	0.633	[3.66]	10.399	[10.74]	0.112	[26.98]
Med	Low	59.458	[69.35]	0.842	[46.78]	13.149	[24.74]	0.274	[40.99]
Low	High	50.452		0.764		11.753		0.190	
Low	Med	52.881	[6.49]	0.786	[4.30]	12.024	[2.16]	0.232	[9.80]
Low	Low	54.945	[4.67]	0.841	[9.00]	12.512	[3.15]	0.272	[6.60]

**Table 4- 9: Does local comovement reflect local information?**

This table reports the results from regressions of twin return differentials on the composite local habitat factor (specification 1 and 2), or separately on the local and foreign habitat factors (specifications 3 to 6). For additional details on the regression specification, see Table 4 and 5. Panel A is reported for a sample where the trading locations of the ETF-pairs and their underlying securities match, while Panel B is reported for a sub-sample where the trading location of ETF-pairs is different from to the trading location of their underlying securities. Coefficients are shown for the average, the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentile of the distribution. T-statistics for the mean are adjusted for cross-correlation as in Hameed, Kang and Viswanathan (2010). T-statistics at the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> refer to fund specific t-statistics and are based on white standard errors. A. R<sup>2</sup> is the average R-squared across all funds, Nobs is the number of observations. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % levels respectively.

	Daily	Monthly	Daily	Monthly		
	1: $\beta_{L-F}$	2: $\beta_{L-F}$	3: $\beta_L$	4: $\beta_F$	5: $\beta_L$	6: $\beta_F$
<b>PANEL A: ETF trading location = underlying stock trading location</b>						
Average coeff.	0.609 (51.82)	0.483 (11.82)	0.623 (27.92)	-0.659 (-40.88)	0.397 (6.13)	-0.672 (-10.58)
Impact (bp)	0.221	0.216	0.117	-0.178	0.115	-0.203
Impact (% SD)	0.484	0.380	0.265	-0.400	0.194	-0.367
Avg. R <sup>2</sup>	0.57	0.49	0.59		0.53	
N obs	273,367	125,08	273,367		12,508	
<b>PANEL B: ETF trading location <math>\neq</math> underlying stock trading location</b>						
Average	0.576 (49.86)	0.469 (8.55)	0.606 (28.94)	-0.610 (-35.46)	0.497 (6.00)	-0.620 (-7.30)
Impact (bp)	0.199	0.206	0.102	-0.170	0.140	-0.207
Impact (% SD)	0.459	0.381	0.265	-0.394	0.265	-0.370
Avg. R <sup>2</sup>	0.54	0.44	0.56		0.49	
N obs	229,033	10,394	229,033		10,394	

**Table 4- 10: Return predictability and reversal of mispricing**

This table reports results for estimating pooled OLS regressions of ETF returns in the local (Panel A) or foreign (Panel B) market on lagged ETF premiums. In order to distinguish between misvaluation in the local and foreign markets, I calculate ETF premiums relative to their underlying portfolio Net Asset Values. The premium of ETF  $i$  in the local market is defined as  $PREM_{i,t} = \ln(P_{i,t} / NAV_{i,t})$ , while the premium of ETF  $i$ 's shares in the foreign market are defined as  $PREM_{p_i,t} = \ln(P_{p_i,t} / NAV_{i,t})$ . All regressions include ETF-pair fixed effects. The t-statistics reported in parentheses beneath the estimates use serial and cross-sectional correlation adjusted Driscoll and Kraay (1998) standard errors with the maximum lag for autocorrelation determined by the number of overlapping days per return horizon. \*/\*\*/\*\* denotes statistical significance at the 10, 5 and 1 % levels respectively.

Variable	Return horizon					
	Day	Week	Month	Quarter	Semi-annum	Year
<b>PANEL A: Y = Local ETF returns</b>						
PREM Local, t-1	-0.149*** (21.56)	-0.071*** (12.54)	-0.042*** (6.72)	-0.025*** (4.49)	-0.017*** (3.23)	-0.012* (1.71)
PREM Foreign, t-1	0.076*** (7.75)	0.029*** (3.69)	0.012 (1.39)	0.010 (1.28)	-0.003 (0.47)	-0.012 (1.29)
R <sup>2</sup>	0.016	0.005	0.007	0.015	0.025	0.055
N obs	498,492	496,893	490,493	474,502	448,515	398,149
<b>PANEL B: Y = Foreign ETF returns</b>						
PREM Local, t-1	0.079*** (11.81)	0.045*** (8.02)	0.018*** (2.63)	0.008 (1.34)	0.005 (0.83)	0.004 (0.64)
PREM Foreign, t-1	-0.134*** (12.77)	-0.079*** (9.78)	-0.044*** (4.67)	-0.024*** (2.83)	-0.024*** (2.97)	-0.027*** (2.91)
R <sup>2</sup>	0.014	0.006	0.007	0.015	0.024	0.053
N obs	498,492	496,893	490,493	474,502	448,515	398,149

## CHAPTER FIVE: CONCLUSION

Exchange-Traded Funds provide investors with easy access to popular investment styles (e.g. large, small, value, growth and sector) and foreign stock market indices at low costs. Apart from being exchange-traded with real-time prices, a unique feature of ETFs is their open-end structure via the share creation/redemption process, which facilitates arbitrage and adds an additional layer of liquidity on top of the liquidity of the ETF in the secondary market.

In the *second chapter* I build on this idea of a liquidity clientele in ETFs. Specifically, I ask whether the liquidity of ETFs relative to its underlying basket encourages shorter-term trading (greater net demand for ETFs at shorter horizons and more institutional buying as well as selling), greater institutional ownership by shorter-term investors relative to longer-term and reduced institutional holding periods. As a starting point, I calculate the relative liquidity for each ETF in the sample. To illustrate, relative liquidity via quoted spreads is highest for small-cap ETFs, followed by mid- and by large-cap. After the financial crisis in 2008, the underlying basket spreads of small-cap ETFs have been around 50 to 100 percent higher than that for the ETF itself. As for Amihud's price impact, I document an increasing trend over time in the large- and small-cap size categories. The average small-cap fund has consistently provided positive relative liquidity via price impact, while for the average large-cap funds relative liquidity is

negative. There are also considerable variations in relative liquidity over time. For ETFs that are less liquid than their underlying portfolios at a given point in time, investors may instead buy ETFs in the primary market at the cost of the underlying securities, or they can time their purchase to coincide with a negative premium.

If liquidity truly facilitates short-term trading, then fund flows at higher frequencies should be more responsive to liquidity than at lower frequencies. I show that relative liquidity (via all four dimensions) indeed predicts net flows positively, particularly for weekly and monthly flows. In contrast, trading costs that are more relevant for longer-term investors – namely expense ratios – are more than 10 times stronger for quarterly net flows as opposed to weekly.

The decisions that I am considering are that of institutional investors, who have the ability to choose between investing in an ETF and its underlying basket. For most retail investors, the transaction costs involved in trading the underlying basket are likely prohibitive. To investigate the hypothesis that relative liquidity matters primarily for institutional investors, I investigate institutional buying and selling separately, defined as the aggregate number of shares bought (sold) by 13-F investors during a quarter scaled by ETF shares outstanding at the beginning of the quarter. Our results demonstrate that ETF liquidity is a strong predictor of institutional buying, and an even stronger predictor institutional selling. In contrast to Cella, Ellul and Giannetti (2013) who find that institutional investors have an incentive to exit during episodes of market turmoil when overall liquidity is already depressed, our results speak to the overall importance of liquidity for selling, not just during episodes of turmoil.

In terms of more direct evidence of a liquidity clientele effect amongst institutional investors, I show that liquidity is more important for investors with short-term trading needs. First, institutional ownership is significantly higher among ETFs with high vs. low relative liquidity (50.81 vs. 33.22 percent institutional ownership). Moreover, institutional investors account for more than 80 percent of the trading in ETFs (Aggarwal and Schofield, 2012), while in our sample their total ownership is on average roughly 40 percent, which suggests that institutions are more active traders in ETFs. I also look separately at institutional ownership by short- and long-term investors and show that relative liquidity is two to six times as important for short- than for long-term institutional ownership (significant at the one percent level).

The incidence of short-term trading in ETFs is also evident from a comparison of holding durations of ETFs relative to their underlying securities. To illustrate, using quarterly holdings

data from SEC 13F filings, I find that the mean holding period for our sample of ETFs is 2.8 quarters, which is on average two quarters *lower* than the dollar-weighted holding duration of its underlying basket. I also compare the fraction of short-term ownership in ETFs, using the definition of transient investors by Brian Bushee (1998), with the fraction of short-term ownership in the underlying securities, and find that ETFs have on average 13.1 percent more short-term investors.

These findings provide direct evidence for a liquidity clientele effect in ETFs. The attractiveness of liquidity for short-term investors may, in the words of John Boogle, Founder of the Vanguard Group, create a temptation to trade. As illustrated in the second chapter, the high liquidity of ETFs attracts noise-traders with correlated demand for investment styles. One important implication of our findings is that as the ETF market continues to grow, it will likely facilitate a higher volume of short-term trading, thereby having an impact on market efficiency and volatility, as well as causing higher correlations in the underlying basket of securities.

In the *third chapter* I argue that, due to the ease of investing in investment styles with ETFs and because of their high liquidity, ETFs attract a clientele of short-term investors with correlated non-fundamental demand for investment styles. In this case, liquidity can sometimes be detrimental for pricing efficiency, because the high liquidity of ETFs attracts short-term investors with correlated non-fundamental demand for investment styles. My findings indicate that there is significant commonality in ETF misvaluation at the investment style level (size, valuation and sector): changes in ETF misvaluation (proxied by the ETF-NAV return) comove positively across ETFs in similar styles, and negatively with ETFs in distant styles. These findings agree with the style switching prediction by Barberis and Shleifer (2003), whereby investors allocate capital at the level of the investment style, and reallocate capital across styles as one style outperforms its distant style. Although the negative distant-style comovements are much weaker in magnitude relative to the own-style comovements, this is not surprising because investors may finance the increase in own-style allocations by selling securities in more than just one (opposite) style. This is particularly true for styles with less well-defined opposites (such as specific sectors, or mid-cap or blend funds).

The own-style misvaluation factor, proxied by the average change in misvaluation of ETFs in a given style, accounts for roughly half of the variation in ETF-NAV returns (regardless of the return horizon used). The economic magnitude, as measured by the impact of a one Std. Dev.

shock to the own-style factor relative to the Std. Dev. of raw ETF returns is considerable at 4 percent at the daily horizon, but declines with the return horizon to 1.27 percent at the monthly horizon. This finding suggests that the proportion of common non-fundamental risk in ETF returns remains roughly constant regardless of the return horizon used, but the economic magnitude (relative to raw returns) declines because arbitrage forces are overall more potent at longer horizon and eliminate both idiosyncratic and common misvaluation.

The evidence of style-based commonality in misvaluation implies excess comovement of returns, but of which, the ETF or the NAV? To provide more direct evidence that changes in ETF misvaluation are in fact driven by misvaluation in the ETF, rather than the NAV leg, I investigate the source of misvaluation. Specifically, if an ETF is hit by a positive *non-fundamental* demand shock that pushes its price above the underlying portfolio NAV value (resulting in a positive ETF premium), then I should observe a reversal in the future ETF returns without any impact on NAV returns. Conversely, if the initial positive demand shock was driven by positive *fundamental* news that is absorbed first into ETF prices, then future NAV returns should be positive as the NAV catches up with a lag and the ETF return should remain unaffected. The empirical results confirm that ETF premiums have a negative and significant impact on future ETF returns over a period of three to four days, consistent with premiums reflecting non-fundamental demand shock. More importantly, reversals in ETF returns are strongest among small-cap ETFs (the category with the highest liquidity relative to their underlying basket), which is consistent with the conjecture that liquidity attract short-term investors with a greater exposure to non-fundamental demand shocks. This finding also rejects the alternative hypothesis that stale pricing among underlying small-cap stocks is responsible for the common factor in misvaluation among small-cap ETFs. Long-term buy-and-hold investors can take advantage of the misvaluation by timing their purchases to coincide with a point in time where premiums are at or below zero.

Moreover, current ETF premiums also forecast future NAV returns negatively over a four day period (positive on first day, negative on the remaining days), which is opposite to what the information diffusion view would predict. Such a negative relationship can, however, arise when investors experience non-fundamental demand shocks and trade sequentially. In this case liquidity goes first to the most liquid securities (ETFs) and when liquidity dries up, demand goes to the next most liquid ETF and so on, until no more ETFs are sufficiently liquid relative to their

underlying securities, in which case the demand goes to the underlying securities. Thus, when ETF premiums reflect large demand shocks, it is likely that both ETFs and their underlying securities are affected, but the impact is on average greater for ETFs due to their high liquidity.

I also provide additional evidence that commonality in misvaluation is in fact driven by demand shocks. First, I find similar style-based comovements in proxies for relative demand shocks (relative turnover and relative ETF price impact) with most of the effect originating from ETF, rather than the NAV leg. Existing studies argue that commonality in liquidity arises either from either from the demand or the supply side. The demand-side view argues that commonality in liquidity arises because of correlated trading activity (Chordia, Roll and Subrahmanyam, 2000; Karolyi, Lee and Van Dijk, 2012), demand by institutional owners (Kamara, Lou and Sadka, 2008), by investor sentiment (Huberman and Halka, 2001) or by the price impact of correlated liquidity needs (Greenwood and Thesmar, 2011). In this case I can view commonality in ETF liquidity as a proxy for correlated demand. The supply-side view argues that liquidity commonality is explained by the funding constraints of financial intermediaries. Several theoretical models predict that commonality in liquidity, via illiquidity spirals or feedback loops, increases during periods when arbitrage capital is limited. My evidence of style-based commonality in liquidity has not been documented in prior literature. This result is more consistent with commonality in demand shocks rather than commonality in the fund constraints of financial intermediaries because the latter is typically characterized as a market-wide (rather than a style-based) phenomenon.

Finally, I find that the magnitude of commonality in misvaluation is significantly stronger for ETFs with greater commonality in demand shocks, and for ETFs with more desirable liquidity characteristics confirming my conjecture that due to the ease of investing in investment styles with ETFs and because of their high liquidity, ETFs attract a clientele of short-term investors with correlated non-fundamental demand for investment styles.

In the *fourth chapter* I investigate whether the trading by a local (country-specific) clientele in Exchange-Traded Funds can induce a country-specific factor in ETF misvaluation. Identification is uniquely achieved by looking at pairs of ETFs from different countries with identical fundamentals (i.e. passively tracking an identical benchmark index), but different clienteles. My findings indicate that twin return differentials (based on local-minus-foreign ETFs), contain a strong country-specific common factor which I refer to as the local habitat

factor. This common factor accounts for almost half of the variation in twin return differentials. The economic significance of the local habitat factors is also strong: a one Std. Dev. increase in the local habitat factor (relative to a one Std. Dev. increase in raw ETF returns) is associated with an increase of 15.34 percent in return differentials at the daily horizon, and a 3.75 percent increase at the monthly horizon. This declining magnitude is consistent with arbitrage activity, but due to limits-to-arbitrage the effect remains highly significant even at the monthly horizon.

An important and interesting alternative explanation is that country-specific commonality in twin return differentials reflects differences in information diffusion between local and foreign securities. Specifically, if local investors are better or faster at processing information about abstract local fundamental factors, then it can give rise to commonality in the twin return differentials of local ETFs relative to their foreign peers. I provide several tests to rule out such an explanation. First, I compare the degree of commonality in misvaluation for two sub-samples: a) a sample where the trading location of ETF-pairs is different from the trading location of their underlying securities with b) a sample where the trading locations of the ETF-pairs and their underlying securities match. The results do not materially differ between the two samples.

Third, I compare the degree of commonality in twin return differentials among ETFs with identical fundamentals, but different levels of liquidity (and non-trading probability). My findings indicate that the degree of commonality in twin return differentials increases from 28.6 to 40.2 to 62.9 percent (as a proportion of total variation in twin return differentials) when the liquidity of the foreign ETF increases from low, to medium to high, holding constant the liquidity of the local ETF. This finding is inconsistent with the alternative explanation of information diffusion, which predicts stronger results for less liquid ETFs that have a higher degree of non-trading (or stale pricing) probability, but the result can be rationalized in the context of local-trading induced misvaluation. In the cross-section of securities, demand is likely to go first to the most liquid securities, implying that commonality in twin return differentials could be stronger for more liquid securities because they attract shorter-term investors. The investors who experience local demand shocks are likely to be shorter-term because in the preferred habitat model, demand shocks reflect changes in sentiment, risk-aversion or liquidity needs, which reflect short-term changes.

If ETF misvaluation truly reflects local-trading induced misvaluation, it should not be confined to the ETF market. Consistent with this hypothesis, I find that changes in ETF

misvaluation are positively (negatively) correlated with stock market returns in the local (foreign) market. This relationship is strongest at the daily horizon, weaker but still significant at weekly horizon, and disappears entirely at the monthly horizon. This finding implies that misvaluation in local stock market returns, at least on a value-weighted basis, declines rapidly with the return horizon. It is certainly possible that specific types of stocks are more exposed to local misvaluation. Specifically, smaller stocks and growth stocks are more likely to be held by local investors (Kumar and Kornioitis, 2013) and may therefore be more exposed to demand shocks by local investors. Although my results reveal a positive a significant relationship between twin return differentials and a local variant of the *SMB* factor, and a negative and significant relationship with *HML* factor at the weekly horizon, this result does not hold at other horizons, nor is it economically very important. In order to properly assess whether local stock returns are also exposed to the local habitat factor (instead of trying to explain changes in ETF misvaluation), I need to formally investigate the asset pricing implications of the local habitat factor for stock returns. I leave this for future research.

## BIBLIOGRAPHY

- Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, 77(2), 375-410.
- Agarwal, S., Liu, C., & Rhee, S. G. (2007). Where does price discovery occur for stocks traded in multiple markets? Evidence from Hong Kong and London. *Journal of International Money and Finance*, 26(1), 46-63.
- Aggarwal, R. and Schofield, L. (2012): *The Growth of Global ETFs and Regulatory Challenges*. Available at SSRN: <http://ssrn.com/abstract=2001060>.
- Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, 5(1), 31-56.
- Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of financial Economics*, 17(2), 223-249.
- Anton, M., & Polk, C. (2014). Connected stocks. *The Journal of Finance*, 69 (3), pp. 1099-1127.
- Atkins, A. B., & Dyl, E. A. (1997). Transactions costs and holding periods for common stocks. *The Journal of Finance*, 52(1), 309-325.
- Bae, K. H., Stulz, R. M., & Tan, H. (2008). Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts. *Journal of Financial Economics*, 88(3), 581-606.
- Baik, B., Kang, J. K., & Kim, J. M. (2010). Local institutional investors, information asymmetries, and equity returns. *Journal of financial economics*, 97(1), 81-106.
- Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, 61(4), 1645-1680.
- Banerjee, A., 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107, 797-817.
- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, 9(1), 3-18.
- Barber, B. M., Odean, T., & Zheng, L. (2005). Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows. *The Journal of Business*, 78(6), 2095-2120.
- Barber, B. M., Odean, T., & Zhu, N. (2009). Do retail trades move markets?. *Review of Financial Studies*, 22(1), 151-186.
- Barberis, N., & Shleifer, A. (2003). Style investing. *Journal of Financial Economics*, 68(2), 161-199.
- Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1, 1053-1128.
- Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, 75(2), 283-317.
- Bartram, S. M., Griffin, J. M., Lim, T. H., & Ng, D. T. (2014). *How Important are Foreign Ownership Linkages for International Stock Returns?*. Available at SSRN 2022129
- Bekaert, G., & Wang, X. (2009). Home bias revisited. Available at: <http://ssrn.com/abstract=1344880>
- Bekaert, G., Hodrick, R. J., & Zhang, X. (2009). International stock return comovements. *The Journal of Finance*, 64(6), 2591-2626.

- Ben-David, I., Franzoni, F., and Moussawi, R. (2014): *Do ETFs Increase Volatility?* Available at SSRN: <http://ssrn.com/abstract=1967599>.
- Bennett, J. A., Sias, R. W., & Starks, L. T. (2003). Greener pastures and the impact of dynamic institutional preferences. *Review of Financial Studies*, 16(4), 1203-1238.
- Ben-Rephael, A., Kadan, O., and Wohl, A. (2013): The Diminishing Liquidity Premium. *Journal of Financial and Quantitative Analysis*, *Forthcoming*.
- Berk, J. B. and Green, R. C. (2004). Mutual fund flows and performance in rational markets, *Journal of Political Economy*, 112(6):1269-1295.
- Bernile, G., Korniotis, G., Kumar, A., & Wang, Q. (2013). Local business cycles and local liquidity. *Journal of Financial and Quantitative Analysis*, *forthcoming*.
- Blackrock (2012): *ETP Landscape, Global handbook 2012*. Available at: <http://www.blackrockinternational.com/intermediaries/en-fi/insights/etp-landscape>
- Blouin, J., Hail, L., & Yetman, M. H. (2009). Capital gains taxes, pricing spreads, and arbitrage: Evidence from cross-listed firms in the US. *The Accounting Review*, 84(5), 1321-1361.
- Bodurtha, J.N., Kim, D-S., Lee, C.M.C (1995): Closed-end Country Funds and U.S. Market Sentiment, *The Review of Financial Studies*, Vol. 8, No. 3, pp. 879-918.
- Boehmer, E. and Kelley, E. (2009): Institutional investors and the informational efficiency of prices, *Review of Financial Studies*, 22, 3563–3594.
- Brav, A., Heaton, J. B., & Li, S. (2010). The limits of the limits of arbitrage. *Review of Finance*, 14(1), 157-187.
- Brockman, P., Chung, D. Y., & Pérignon, C. (2009). Commonality in liquidity: A global perspective. *Journal of Financial and Quantitative Analysis*, 44(04), 851-882.
- Broman, M. S. (2015). *Liquidity, Style investing and Excess Comovement of Exchange-Traded Fund Returns*. Working paper, York University.
- Broman, M. S. and Shum, P. (2015): *Short-Term Trading and Liquidity Clienteles: Evidence from Exchange-Traded Funds*. Working paper, York University.
- Broman, M. S. and Shum, P. (2015b): *Positive feedback style trading in ETFs*. York University unpublished working paper.
- Brown, S. J., & Goetzmann, W. N. (1997). Mutual fund styles. *Journal of financial Economics*, 43(3), 373-399.
- Bushee, B. J. (1998). Investors on Myopic R&D Investment Behavior. *The accounting review*.
- Campbell, J. Y., Giglio, S., Polk, C., & Turley, R. (2013). *An Intertemporal CAPM with Stochastic Volatility* (No. w18411). National Bureau of Economic Research.
- Campbell, J. Y., Polk, C., & Vuolteenaho, T. (2010). Growth or glamour? Fundamentals and systematic risk in stock returns. *Review of Financial Studies*, 23(1), 305-344.
- Campbell, J. Y., Ramadorai, T., & Schwartz, A. (2009). Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics*, 92(1), 66-91.
- Chan, J. S., Hong, D., & Subrahmanyam, M. G. (2008). A tale of two prices: Liquidity and asset prices in multiple markets. *Journal of Banking & Finance*, 32(6), 947-960.

- Chan, K., Covrig, V., & Ng, L. (2005). What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide. *The Journal of Finance*, 60(3), 1495-1534.
- Chan, K., Hameed, A., & Ting Lau, S. (2003). What if trading location is different from business location? Evidence from the Jardine Group. *The Journal of Finance*, 58(3), 1221-1246.
- Chan, L. K., Chen, H. L., & Lakonishok, J. (2002). On mutual fund investment styles. *Review of financial studies*, 15(5), 1407-1437.
- Chen, H., & Kim, H. (2008). American depositary receipts: Asia–Pacific evidence on convergence and dynamics. *Journal of Multinational Financial Management*, 18(4), 346-368.
- Cherkes, M., Sagi, J., & Stanton, R. (2009). A liquidity-based theory of closed-end funds. *Review of Financial Studies*, 22(1), 257-297.
- Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics*, 56(1), 3-28.
- Clifford, C.P., Fulkerson, J.A. and Jordan, B.D. (2014): What drives ETF flows? *The Financial Review*, 49, pp. 619-642.
- Cooper, M. J., Gulen, H., & Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. *The Journal of Finance*, 60(6), 2825-2858.
- Coval, J. D., & Moskowitz, T. J. (2001). The geography of investment: Informed trading and asset prices. *Journal of Political Economy*, 109(4), 811-841.
- Coval, J., & Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, 86(2), 479-512.
- Cremers, M., & Pareek, A. (2014): Short-Term Trading and Stock Return Anomalies: Momentum, Reversal, and Share Issuance. *Review of Finance*, forthcoming.
- Da, Z., & Shive, S. (2013): *When the Bellwether Dances to Noise: Evidence from Exchange-Traded Funds*. Available at SSRN: <http://ssrn.com/abstract=2158361>.
- Daniel, K., D. Hirshleifer, and A. Subrahmanyam. 2001. Overconfidence, Arbitrage, and Equi-librium Asset Pricing. *Journal of Finance*, 56: 921-965.
- DeLong, J. B., Shleifer, A., Summers, L.H. and Waldmann, R. (1990): Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703-738.
- Deutsche Bank (2011): *ETF Research – Industry Perspective - In the ETF labyrinth, where does the thread begin?* Deutsche Bank (7th July).
- Driscoll, John C., and Aart C. Kraay (1998): Consistent covariance matrix estimation with spatially dependent panel data, *Review of Economics and Statistics* 80, pp. 549–560.
- Edelen, R. M. (1999). Investor flows and the assessed performance of open-end mutual funds. *Journal of Financial Economics*, 53(3), 439-466.
- Edelen, R. M., & Warner, J. B. (2001). Aggregate price effects of institutional trading: a study of mutual fund flow and market returns. *Journal of Financial Economics*, 59(2), 195-220.
- Engle, R. and Sarkar, D. (2006): Premiums-Discounts and Exchange Traded Funds. *The Journal of Derivatives*, 2006 (1), pp. 27-45.

- Eun, C. S., & Sabherwal, S. (2003). Cross - Border Listings and Price Discovery: Evidence from US - Listed Canadian Stocks. *The Journal of Finance*, 58(2), pp. 549-576.
- Fama, E. F., & French, K. R. (2012): Size, value, and momentum in international stock returns. *Journal of financial economics*, 105(3), 457-472.
- Fama, Eugene F., and Kenneth R. French, (1993): Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.
- Fama, Eugene F., and Kenneth R. French, (1995): Size and book-to-market factors in earnings and returns, *Journal of Finance*, 50, 131–155.
- Feng, L., & Seasholes, M. S. (2004): Correlated trading and location. *The Journal of Finance*, 59(5), 2117-2144.
- Ferreira, M. A., & Matos, P. (2008): The colors of investors' money: The role of institutional investors around the world. *Journal of Financial Economics*, 88(3), 499-533.
- Fong, K., Holden, C.W., Trzcinka, C.A. (2010): *Can global stock liquidity be measured?* Unpublished working paper. University of New South Wales and Indiana University.
- Frazzini, A., & Lamont, O. A. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics*, 88(2), 299-322.
- Froot, K., & Teo, M. (2008). Style investing and institutional investors. *Journal of Financial and Quantitative Analysis*, 43(4), 883.
- Froot, K., Scharfstein, D., Stein, J., 1992. Herd on the street: informational inefficiencies in a market with short-term speculation. *Journal of Finance* 47, 1461–1484.
- Gagnon, L., & Andrew Karolyi, G. (2010). Multi-market trading and arbitrage. *Journal of Financial Economics*, 97(1), 53-80.
- Gompers, P. A., & Metrick, A. (2001). Institutional investors and equity prices. *The quarterly journal of Economics*, 116(1), 229-259.
- Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity?. *Journal of Financial Economics*, 92(2), 153-181.
- Graham, J. R., Li, S., & Qiu, J. (2012). Managerial attributes and executive compensation. *Review of Financial Studies*, 25(1), 144-186.
- Green, T. C., & Hwang, B. H. (2009). Price-based return comovement. *Journal of Financial Economics*, 93(1), 37-50.
- Greenwich Associates (2013): *Institutional Investors Relation with ETFs deepen*. Available at: [http://us.ishares.com/content/en\\_us/repository/resource/greenwich\\_survey\\_inst.pdf](http://us.ishares.com/content/en_us/repository/resource/greenwich_survey_inst.pdf)
- Greenwood, R. (2007): Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. *Review of Financial Studies*, 21.3, pp. 1153-1186.
- Greenwood, R., & Thesmar, D. (2011). Stock price fragility. *Journal of Financial Economics*, 102(3), 471-490.
- Grinblatt, M., Ikäheimo, S., Keloharju, M., & Knüpfer, S. (2013). IQ and mutual fund choice. Available at SSRN: <http://ssrn.com/abstract=2021957>
- Gruber, M., 1996. Another puzzle: the growth in actively managed mutual funds. *Journal of Finance*, 51, 783–810.

- Hameed, A., Kang, W., & Viswanathan, S. (2010). Stock market declines and liquidity. *The Journal of Finance*, 65(1), 257-293.
- Hardouvelis, G., Porta, R. L., & Wizman, T. A. (1994). *What moves the discount on country equity funds?*. In *The internationalization of equity markets* (pp. 345-403). University of Chicago Press.
- Hasbrouck, J. (2009). Trading costs and returns for US equities: Estimating effective costs from daily data. *The Journal of Finance*, 64(3), 1445-1477.
- Hirshleifer, D., Subrahmanyam, A., Titman, S., 1994. Security analysis and trading patterns when some investors receive information before others. *Journal of Finance* 49, 1665–1698.
- Hu, G. X., Pan, J., & Wang, J. (2013). Noise as information for illiquidity. *The Journal of Finance*, 68(6), 2341-2382.
- Huang, M. (2003). Liquidity shocks and equilibrium liquidity premia. *Journal of Economic Theory*, 109(1), 104-129.
- Huberman, G., Halka, D. (2001): Systematic liquidity. *Journal of Financial Research*, 24, pp. 161–178.
- Idzorek, T. M, Xiong, J. X. and Ibbotson, R. G. (2012). The Liquidity Style of Mutual Funds. *Financial Analyst Journal*, 68 (6),
- Jain, P. C., & Wu, J. S. (2000). Truth in mutual fund advertising: Evidence on future performance and fund flows. *The Journal of Finance*, 55(2), 937-958.
- Kamara, A., Lou, X., & Sadka, R. (2008). The divergence of liquidity commonality in the cross-section of stocks. *Journal of Financial Economics*, 89(3), 444-466.
- Karolyi, G. A., Lee, K. H., & Van Dijk, M. A. (2012). Understanding commonality in liquidity around the world. *Journal of Financial Economics*, 105(1), 82-112.
- Kasch, M., & Sarkar, A. (2012). Is there an S&P 500 Index effect?. Available at SSRN 2171235.
- Ke, D., Ng, L., & Wang, Q. (2010). Home bias in foreign investment decisions. *Journal of International Business Studies*, 41(6), 960-979.
- Korajczyk, R. A., & Sadka, R. (2008). Pricing the commonality across alternative measures of liquidity. *Journal of Financial Economics*, 87(1), 45-72.
- Korniotis, George M. (2008): Habit formation, incomplete markets, and the significance of regional risk for expected returns, *Review of Financial Studies*, 21, pp. 2139–2172.
- Korniotis, G. M., & Kumar, A. (2013). State - Level Business Cycles and Local Return Predictability. *The Journal of Finance*, 68(3), 1037-1096.
- Kumar, A. (2009). Dynamic style preferences of individual investors and stock returns. *Journal of Financial and Quantitative Analysis*, 44(03), 607-640.
- Kumar, A. (2009). Hard-to-value stocks, behavioral biases, and informed trading. *Journal of Financial and Quantitative Analysis*, 44(06), 1375-1401.
- Kumar, A., & Lee, C. (2006). Retail investor sentiment and return comovements. *The Journal of Finance*, 61(5), 2451-2486.
- Kumar, A., Page, J. K., & Spalt, O. G. (2013). Investor Sentiment and Return Comovements: Evidence from Stock Splits and Headquarters Changes. *Review of Finance*, 17(3), 921-953.

- Kumar, A., Page, J. K., & Spalt, O. G. (2013b): Gambling and Comovement. *Journal of Financial and Quantitative Analysis*, forthcoming.
- Lee, C. M. C., Shleifer A., & Thaler, R. (1990). Anomalies: Closed-End Mutual Funds. *Journal of Economic Perspectives*, 4(04), 153-164.
- Lee, C.M., Shleifer, A. and Thaler, R.H. (1991): Investor Sentiment and the Closed-End Fund Puzzle. *The Journal of Finance*, XLVI (1), pp. 75-109.
- Lee, K. H. (2011). The world price of liquidity risk. *Journal of Financial Economics*, 99(1), 136-161.
- Lesmond, D. (2005): Liquidity of emerging markets. *Journal of Financial Economics*, 77, 411–452.
- Lo, A. W., & Wang, J. (2000). Trading volume: definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies*, 13(2), 257-300.
- Lo, A., W., Mamaysky, H., & Wang, J. (2004). Asset Prices and Trading Volume under Fixed Transactions Costs. *Journal of Political Economy*, 112(5).
- Lou, D. (2012). A flow-based explanation for return predictability. *Review of Financial Studies*, 25(12), 3457-3489.
- Lynch, A. W., & Tan, S. (2011). Explaining the Magnitude of Liquidity Premia: The Roles of Return Predictability, Wealth Shocks, and State-Dependent Transaction Costs. *The Journal of Finance*, 66(4), 1329-1368.
- Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2013). ETF arbitrage: Intraday Evidence. *Journal of Banking and Finance*, 37(9), 3486-3498.
- Merton, R. 1987. Presidential Address: A Simple Model of Capital Market Equilibrium with Incomplete Information. *Journal of Finance*, 42:483–510.
- Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867-887.
- Morck, R., Yeung, B., & Yu, W. (2000): The information content of stock markets: why do emerging markets have synchronous stock price movements?. *Journal of financial economics*, 58(1), 215-260.
- Parwada, J., & Siaw, K. (2013). Empirical Test of the Liquidity-Based Theory of Closed-End Funds. Available at SSRN 2267599.
- Pástor, L., & Stambaugh, R. F. (2003). Liquidity Risk and Expected Stock Returns. *The Journal of Political Economy*, 111, 642-685.
- Peng, L. and W. Xiong, 2006. Investor Attention, Overconfidence and Category Learning. *Journal of Financial Economics* 80: 563-602.
- Perez, F., Shkilko, A., & Tang, T. (2011): *Signaling Via Stock Splits: Evidence from Short Interest*. In AFA 2012 Chicago Meetings Paper.
- Petajisto, A (2013): *Inefficiencies in the Pricing of Exchange-Traded Funds*. Available at SSRN: <http://ssrn.com/abstract=2000336>.
- Pirinsky, C., & Wang, Q. (2006). Does corporate headquarters location matter for stock returns? *The Journal of Finance*, 61(4), 1991-2015.
- Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics*, 42(1), 35-52.

- Richie, Nivine, Robert Daigler, and Kimberly C. Gleason (2008): The limits to stock index arbitrage: Examining S&P 500 futures and SPDRs, *Journal of Futures Markets*, 28(12), 1182– 1205.
- Schmidt, P. S., Von Arx, U., Schrimpf, A., Wagner, A. F., & Ziegler, A. (2014). *On the construction of common size, value and momentum factors in international stock markets: A guide with applications*. Swiss Finance Institute Research Paper, (10-58).
- Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *The Journal of Economic Perspectives*, 19-33.
- Shum, P. & Chen, K. Y. (2012). All are NOT Created Equal: An Exploration of ETF Asset Growth, *Journal of Index Investing*, vol.3-3, 49-61.
- Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *The Journal of finance*, 53(5), 1589-1622.
- Staer, A. (2014). *Fund Flows and Underlying Returns: The Case of ETFs*. <http://ssrn.com/abstract=2158468>
- Teo, M. (2009). The Geography of Hedge Funds. *Review of Financial Studies*, 22(9), 3531-3561.
- Teo, M., & Woo, S. J. (2004). Style effects in the cross-section of stock returns. *Journal of Financial Economics*, 74(2), 367-398.
- Tversky, A., Kahneman, D. (1974) Judgement under uncertainty: heuristics and biases. *Science*, 185, 1124–1131.
- Van Nieuwerburgh, S., & Veldkamp, L. (2009). Information immobility and the home bias puzzle. *The Journal of Finance*, 64(3), 1187-1215.
- Vayanos, D., and J. Wang, 2012, Liquidity and asset returns under asymmetric information and imperfect competition, *Review of Financial Studies*, 25, 1339-1365
- Wurgler, Jeffrey (2010): *On the Economic Consequences of Index-Linked Investing*, National Bureau of Economic Research Working Paper Series No. 16376.
- Yan, X. S. (2008). Liquidity, investment style, and the relation between fund size and fund performance. *Journal of Financial and Quantitative Analysis*, 43(03), 741-767.
- Ye, P. (2012). The value of active investing: can active institutional investors remove excess comovement of stock returns?. *Journal of Financial and Quantitative Analysis*, 47(03), 667-688.
- Zhang, K. H. C. H., & Chung, K. H. (2009). A simple approximation of intraday spreads using daily data. *Journal of Financial Markets*, pp 94-120.
- Zheng, L., 1999. Is money smart? A study of mutual fund investors' fund selection ability. *Journal of Finance*, 54, 901–933.

## APPENDICES

### APPENDIX A3-1: VARIABLE DESCRIPTIONS

Variables	Description / Calculation	Source
$IFL_t$	Dollar value of daily net share creations over the period $t$ (week, month or quarter) scaled by beginning-of-period AUM.	Bloomberg
$CIFL_{t-4}$	$\ln(1 + \text{cumulative dollar inflows up to month } t-4 / \text{AUM on 01/2006})$	Bloomberg
$ETF\ Ret$	Monthly ETF return inclusive of dividends	CRSP
$AUM$	Number of shares outstanding multiplied by NAV.	Bloomberg
$Age$	$(\text{Current Date-Inception})/365$	Morningstar
$PREM$	Average monthly $PREM$ , where $PREM_{d-1} = \ln(\text{ETF mid-price}) - \ln(\text{NAV})$	CRSP / Bloomberg
$Track. Err.$	Tracking error, i.e. the monthly standard deviation of the daily change in premiums.	CRSP / Bloomberg
$EXP\ Ratio$	The percentage of fund assets used to pay for operating expenses and management fees, including 12b-1 fees, administrative fees, and all other asset-based costs incurred by the fund, except brokerage costs. Sales charges are not included.	Morningstar
$VIX$	S&P 500 option implied volatility	CBOE
$TS$	The difference in yields on a 10 year U.S. government bond yield and a 3 month U.S. T-bill	Datastream
$P\&S\ LIQ$	Pastor and Stambaugh (2003) measure of aggregate liquidity	WRDS
<b>Liquidity measures (Note: all variables are signed such that higher values indicate greater liquidity)</b>		
Quoted Spread	$QSPR_{i,t} = \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \frac{ASK_{i,d} - BID_{i,d}}{(ASK_{i,d} + BID_{i,d})/2} \right]$ <p>where <math>ASK</math> and <math>BID</math> refer to closing <math>ASK</math> and <math>BID</math> prices of the ETF <math>i</math> on day <math>d</math>. <math>N_t</math> refers to the number of days in month <math>t</math>.</p>	CRSP
Price Impact	$AMIHU_{i,t} = \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \frac{ R_{i,d} }{DVOL_{i,d}} \right]$ <p>where <math>R</math> is the ETF return (in %) and <math>DVOL</math> is the dollar volume on day <math>d</math>. Daily observations beyond the 99.5<sup>th</sup> percentile are discarded.</p>	CRSP
Turnover	$TO_{i,t} = \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \frac{VOL_{i,d}}{SHR_{i,d}} \right]$ <p>where <math>VOL</math> is the share volume of ETF <math>i</math> on day <math>d</math>, <math>SHR</math> is shares outstanding. Daily observations beyond the 99.5<sup>th</sup> percentile are discarded.</p>	CRSP / Bloomberg
Share Creation Activity	$CREATE_{i,t} = \log \left[ 1 + \frac{1}{N_t} \sum_{d=1}^{N_t} \frac{ SHR_{i,d} - SHR_{i,d-1} }{SHR_{i,d-1}} \right]$ <p>where <math>SHR</math> is shares outstanding for ETF <math>i</math> on day <math>d</math> (from Bloomberg)</p> <p>Relative liquidity is the ratio of ETF to underlying portfolio liquidity:</p>	Bloomberg
REL(LIQ)	$REL(LIQ_{i,t}) = \frac{1}{N_t} \sum_{d=1}^{N_t} \log \left[ \sum_{k=1}^K w_{i,k,d} LIQ_{k,d}^{UND} / LIQ_{i,d}^{ETF} \right]$ <p>where <math>LIQ</math> is one of the three liquidity measures (for Amihud I use the difference as it is in logs). Portfolio liquidity of ETF <math>i</math> is measured as the dollar-weighted average of individual asset <math>k</math>'s (monthly) liquidity.</p>	CRSP / Bloomberg

## **APPENDIX A4-1: ETF-DATA CLEANING**

In constructing the levels and changes in premiums, I take great care to deal with outliers and data errors in the raw data for the ETF and NAV prices from Bloomberg. First, when the ETF mid-quote price is missing, I use the closing price instead. Second, I remove all ETFs from the sample with more than 20 percent missing (non-trading) observations for both the ETF and NAV prices. Also, for ETF-months where more than one third of ETF or NAV prices are missing due to non-trading, I discard that fund-month from the sample. Third, the underlying currency for the ETF and NAV prices is sometimes misreported in Bloomberg resulting in a sequence of extreme positive or extreme negative premiums. In such cases the correct exchange rate is manually applied. Fourth, the time-stamp for the official NAV price is sometimes off by one day compared with the official NAV observations from Morningstar or the iNAV series from Bloomberg (in 1.5 % of daily observations). In such cases I prioritize official NAV prices from Morningstar (when correct) over iNAV prices from Bloomberg. To determine whether the alternative NAV prices have a correct time-stamp, I require that the correlation between the “correct” NAV return and the ETF return is more than 0.5 greater than the correlation between the “wrong” NAV returns and the ETF return.

Next, I take a closer look at the premiums. For a handful of observations the premium based on mid-point prices is far larger than the premium based on closing prices (10 percentage point difference). In such cases I use the premium (and ETF return) based on closing prices instead. A handful of premiums are greater than 20 % in absolute magnitude. These are considered to be data errors similar to Petajisto (2013), and are discarded from the sample.

A simple correlation analysis between ETF, NAV and underlying index returns (UND) can be used to detect anomalous observations. There are three correlations to consider. First, the correlation between the NAV and the UND returns must be close to 1, unless the fund uses sample replication (rare in my sample). Second, the correlation between the NAV and ETF returns should generally also be highly positive, but it can be lower when the ETF is highly mispriced. Third, the correlation between the ETF and the UND return should also be high, but generally lower than that between the ETF and the NAV. I take action when NAV returns (based on iNAVs) are more highly correlated with both ETF and UND returns (by 0.1 on average) relative to official NAV returns, and when the correlation between ETF and UND returns is greater than the average correlation between official NAV and ETF returns, and that between

official NAV and UND returns. In such cases I replace all variables using official NAV prices with those using iNAV prices. Roughly 3 % of observations are affected. Also, when premiums based on official NAV prices are missing, I use iNAV prices and returns instead. To ensure accuracy of iNAV returns, I require that the correlation between ETF and iNAV returns, and that between UND and iNAV returns to be greater than 0.5.

Finally, to mitigate the impact of any remaining outliers, I winsorize all measures of returns and misvaluation at 5 Std. Dev. from the mean fund-by-fund.

## **APPENDIX A4-2: STOCK-DATA CLEANING**

The following static screens are applied to stock-level data from Datastream: I keep major listings (MAJOR="Y"), stocks located in the domestic market (e.g. GEOGN="GERMANY) and firms of the equity type (TYPE="EQ"). I cross-check the security type from Bloomberg to ensure that only common equity is retained (SECURITY\_TYP="Common Stock") and that the primary security of a company trades in the domestic market (e.g. EQY\_PRIM\_SECURITY\_COMP\_EXCH="GR" for Germany). Financial firms are removed from the sample (ICB-sector code = 8000).

The following dynamic screens are employed:

1. In order to minimize potential biases arising from low-price and illiquid stocks, I delete "penny" stocks from the sample. Specifically, stocks with beginning-of-period prices below 0.50 USD. (Schmidt *et al.*, 2014)
2. Discard all price observations where the price is greater than 1,000,000 of the domestic currency. (Schmidt *et al.*, 2014)
3. I set the daily return to be missing if any daily return above 150% (inclusive) is reversed the following day. Specifically, the daily returns for both days  $t$  and  $t-1$  are set to missing if  $[1+r(i,t)][1+r(i,t-1)]-1 < 50\%$ , where  $r(i,t)$ , is the simple return for day  $t$ , and at least one of the two returns is 150% or greater. (Schmidt *et al.*, 2014; Lee, 2011)
4. If daily returns are greater than 375 percent, I exclude returns from day  $-1$  to  $+1$  around the extreme event. (Bartram, Griffin, Lim, and Ng, 2014)
5. To handle splits, mergers, and potential data errors, daily returns of the extreme 0.1% (inclusive) at the top or bottom of the cross-section of each country are set to missing for that month (Amihud, 2002; Hou, Karolyi, and Kho, 2006; Lee, 2011).

### APPENDIX A4-3: THE MECHANICS OF ARBITRAGE AND CROSS-LISTINGS WITH EUROPEAN ETF

ETFs have an open-ended structure that is available to select institutional investors (known as Authorized Participants, APs) via the creation/redemption process. This process is an arbitrage mechanism by which Authorized Participants (APs) exchange the underlying basket for ETF shares with the ETF manager when there is excess demand for the ETF, or when it is trading at a premium relative to Net Asset Value (share creation). APs exchange ETF shares for the underlying basket when the opposite occurs (share redemption).

This arbitrage process is the same for ETFs traded in the U.S. or Europe. The main difference is that in the U.S., ETFs are traded across multiple trading venues (e.g. NYSE ARCA, NASDAQ, BATS, Direct Edge-X), but trades are settled in one depository system (the Depository Trust Company). In contrast, European ETFs are listed on multiple national markets – up to five – and secondary market trades are settled in the central securities depository (CSD) of the country in which the trade was executed. This structure of multiple cross-listings and settlement systems imposes additional costs on arbitrageurs, thereby worsening pricing efficiency.

European ETFs can be listed in one of three ways: 1) the classic-listed model, 2) the Depository Receipt model or 3) the split structure<sup>55</sup>.

First, the *classic-listed model* is typically used by ETFs domiciled in Ireland or Luxembourg<sup>56</sup>, but listed in France. In this case, the ETF sponsor only uses one central securities depository (CSD) to create or redeem ETF shares, referred to as the issuer depository. Other depositories wanting to settle ETF trades are called investor depositories. Cross-border arbitrageurs wanting to transfer newly created ETF shares from the issuer depository to investor depositories (e.g. France) can do so electronically, but this creates additional costs every time shares are moved across depositories<sup>57</sup>. This method cannot be used when depositories are not linked to each other, or if regulators prohibit it (such as in Germany). In some cases, transfers between two markets need to through a third depository (e.g. in order to move an Irish ETF from the Dutch to the Italian market, you would need to go through Germany).

---

<sup>55</sup> <http://www.securitiestechologymonitor.com/blogs/european-etfs-dozens-depositories-dtc-28241-1.html>

<sup>56</sup> As of March 2013, more than 84 % of all European ETFs are domiciled in either Luxembourg, or Ireland (80 % for my sample of countries). <http://www.pwc.lu/en/fund-distribution/docs/pwc-publ-european-etf.pdf>

<sup>57</sup> Financial Times Alphaville posted by Izabella Kaminska on Tuesday September 20th, 2011 13:19

Second, the *depository receipt model* is similar to the American Depository Receipt (ADR) model for cross-listed stocks. This model is often used by ETFs domiciled in Ireland, but listed in Germany. In this case the investor depository is not allowed to transfer shares electronically. Instead, another party (typically a bank) has to re-issue a number of ETFs into the remote depository, lock them away, and issue depository receipts to local investors. This can be costly because ETF arbitrage involves frequent share creations / redemptions, unlike for cross-listed shares. In addition, the process can be operationally challenging as the re-issued ETF receipts may have different identification codes (ISIN's) making it harder for a broker's back-office system to handle and increasing the risk for settlement failures.

Third, the *split structure* is a hybrid of the other two. In this case ETFs are simultaneously issued into multiple depositories (e.g. ETFs domiciled in Ireland are listed in both Germany and the U.K.). This process is operationally complicated because it requires manual processing. A broker-dealer has to send an instruction – often by fax – to each depository to deliver ETFs to a transfer agent in one market, and to receive ETFs from a transfer agent in the other market. This process can take from one to several days, especially if correct communication protocols are not followed, increasing the risk of settlement failure.

These complications in the transfer and settlement of ETF shares - including the requirement to have depository accounts in each country and follow post-trading rules in each jurisdictions - increase operational costs and increase the risk of costly settlement failures, particularly for cross-border arbitrageurs<sup>58</sup>. In an effort to avoid settlement failures, market makers typically deposit their inventories of ETFs at different CSDs across Europe, which leads to a fragmentation of liquidity<sup>59</sup>. This has also been cited as a deterrent for institutional investors and for asset growth<sup>60</sup>. Another problem with the multiple listing & settlement is that prime brokers and clearinghouses may use a so-called omnibus account at the CSD. This means that

---

<sup>58</sup> Managing the settlement procedure efficiently is essential for APs, according Mr Butler of Northern Trust. Financial Times by Chris Flood on September 19, 2011 12:36 am.: <http://www.ft.com/cms/s/0/2cef77ae-d00b-11e0-81e2-00144feabdc0.html>.

European exchanges have a policy of forcing a buy-in, in which case the exchange buys-in the securities at current market value and forcing the defaulting seller to cover the costs. (<http://www.etf.com/publications/journalofindexes/joi-articles/7260-the-fragmentation-of-the-european-etf-market.html>)

<sup>59</sup> “ETFs: building liquidity” on 02 Nov 2010: <http://www.professionaladviser.com/etfm/feature/1866073/building-liquidity>

<sup>60</sup> Financial Times by Chris Flood on February 11, 2014 6:35 pm: <http://www.ft.com/cms/s/0/e382016a-932e-11e3-b07c-00144feab7de.html>

ETFs can be loaned out by the clearinghouse to another investor without the knowledge or approval from the owner of the ETF. ETFs that are lent cannot be transferred, and can therefore not be used in cross-border arbitrage trades.