Journal Pre-proofs

The Price of International Equity ETFs: The Role of Relative Liquidity

Christina Atanasova, Jean-Philippe Weisskopf

PII: S1042-4431(20)30074-3
DOI: https://doi.org/10.1016/j.intfin.2020.101190
Reference: INTFIN 101190

To appear in: *Journal of International Financial Markets, Institutions & Money*

Received Date: 5 November 2019
Accepted Date: 28 February 2020

Please cite this article as: C. Atanasova, J-P. Weisskopf, The Price of International Equity ETFs: The Role of Relative Liquidity, *Journal of International Financial Markets, Institutions & Money* (2020), doi: https://doi.org/10.1016/j.intfin.2020.101190

The Price of International Equity ETFs: The Role of Relative Liquidity *

Christina Atanasova† and Jean-Philippe Weisskopf‡

November 2019

Abstract

We examine the effect of the relative liquidity of international equity exchange-traded funds (ETFs) and their constituent portfolios on the price difference between the fund's market prices and its net asset values. We use data for a sample of 584 international equity ETFs listed in the U.S. over the period January 2012 to December 2017 and find that higher liquidity is associated with a lower absolute value of the ETF premium/discount. We document a positive relationship between liquidity and the price convergence of the ETFs and their underlying shares. The effect of liquidity on convergence is stronger for ETFs with high holding costs.

JEL Classification: G14, G15

Keywords: International equity; exchange-traded fund; premium / discount; liquidity.

*This paper was presented at the 2018 Cross Country Perspectives of Finance conferences held in Guangzhou, China and Dar es Salaam, Tanzania. We thank two anonymous referees, the editor and the participants at both conferences for helpful comments and suggestions.

†Atanasova: Beedie School of Business, Simon Fraser University. E-mail: cva3@sfu.ca.

‡Weisskopf: Ecole hôtelière de Lausanne, HES-SO Haute école spécialisée de Suisse occidentale. Email: jean-philippe.weisskopf@ehl.ch

1 Introduction

International equity exchange-traded funds (ETFs) have grown in popularity as they provide an easy and cheap form of international diversification. An ETF's value is based on the value of the portfolio of securities in the fund and (unlike closed-end funds) an arbitrage mechanism exists where a trader can take long/short positions in the ETF and simultaneously redeem/create shares.[1] Nevertheless, the market value of an ETF can deviate from the net asset value (NAV) of the underlying portfolio of securities. This prevalence of an ETF premium or discount is a well-documented phenomenon. In particular, Petajisto (2017) examines the deviation of midquote prices of exchange-traded funds from their NAVs and shows that it is significantly larger for international funds and funds with illiquid underlying securities.

In addition, ETFs with nearly identical underlying portfolios often trade at values on the opposite side of their NAVs. Two emerging markets funds, tickers EEM (iShares MSCI Emerging Markets Index Fund) and SCHE (Schwab Emerging Markets ETF), offer exposures to very similar well-diversified portfolios of emerging market stocks but, as shown in Figure 1 (Panel A) one sometimes traded at a premium relative to its net asset value while the other traded at a discount at exactly the same time. Developed markets ETFs exhibit similar behavior. Figure 1 (Panel B) shows the ETF premium/discount for EFA (iShares MSCI EAFE Index Fund) and SCHF (Schwab International equity ETF) over the same period. Therefore, it is important for an investor to correctly identify which ETF to trade on a given day.

Previous studies have linked the efficiency of ETF prices to trading costs and other limits to arbitrage that might deter traders profiting from a mispricing. Gastineau (2010) and Tucker and Laipply (2010) argue that the officially published NAV may not fully reflect the current value of the ETF portfolio due to stale pricing. NAV is computed based on the latest closing prices for the underlying securities thus creating a problem for stocks traded in illiquid international markets. Furthermore, in the

[1] There is no such analogue for an open-end mutual fund for which all transactions occur at NAV at the end of the trading day.

case of stocks traded in different time zones it may not be possible to enter into simultaneous offsetting transactions.

Using data for 584 U.S.-listed International Equity ETFs for the period January 3, 2012 to December 29, 2017, this paper analyzes how the ETF premium or discount is related to the relative liquidity between the fund and the portfolio of underlying securities. Most of the literature that links liquidity and the ETF premium/discount has focused on funds that track U.S. equity and bond indices. Dannhauser (2017), for example, studies corporate bond ETFs and finds that higher ETF ownership lowers bond yields but has an insignificant or negative impact on bond liquidity. Chacko, Das and Fan (2016) use ETF premiums as a metric for bond market illiquidity. Broman and Shum (2018) document a liquidity clientele for ETFs that track U.S. equity indices, where liquidity is significantly more important for short- than for long-term investors. Hamm (2010) analyzes the link between the ETF's liquidity and the liquidity of its constituents, while Richie and Madura (2007) examine the impact of the introduction of new ETFs on the underlying stock liquidity.

Bhattacharya and O'Hara (2018) consider the information linkage between ETFs and the underlying portfolio in the case of "hard-to-trade" assets. They show that such ETFs can affect the informational efficiency of the underlying markets and introduce fragility via herding. Malamud (2016) develops a theoretical model that links the creation/redemption of ETFs and how such liquidity shocks propagate to the underlying securities. Ben-David, Franzoni and Moussawi (2018) provide empirical evidence that higher ETF ownership leads to higher intraday and daily volatility of the funds' underlying assets. They show that excess volatility may arise due to staleness in NAV and price discovery in the ETF. They exploit exogenous changes in index membership and find that ETF ownership increases the negative autocorrelation in stock prices and that stocks with higher ETF ownership display significantly higher volatility. Finally, Cong and Xu (2016) explore the optimal design of the ETF baskets.

We deviate from the previous literature by focusing on exchange-traded funds based on international publicly traded equity. U.S. financial markets are characterized by low transaction costs and high levels of liquidity and efficiency. Our paper examines what liquidity actually means for international ETFs and thinly traded markets. We concentrate on international equity as an asset class for several reasons. First and most important, international bonds, shares in privately held companies and derivatives such as warrants, and rights are not traded in public markets but are either traded on illiquid OTC markets or not traded at all. Commodity and currency contracts, on the other hand, are traded globally rather than only on a U.S. exchange (the ETF) and non-U.S. exchanges (the underlying assets). Second, most of the international ETFs traded in the U.S. are equity funds. Table A1, in the appendix, presents the distribution of funds, assets under management (AUM) and fund flows of all ETFs in the ETF Global database by year, asset class and region. For ETFs on non-U.S. assets, international equity funds represent almost 80% of the total number of ETFs and 90% of the assets under management.

Figure 2 illustrates the idea of our paper in a simple way. Figure 2 Panel A presents a time series plot of the median ETF premium/discount and the median ETF illiquidity, whereas Panel B presents a time series plot of the median ETF premium/discount and the median illiquidity of the constituent portfolios. The figure suggests that large absolute values of the ETF premium are associated with periods of high illiquidity.

Our regression analysis examines the relationship between the ETF premium/discount and the liquidity of the ETF and the underlying portfolio of securities. Our first set of results shows that larger differences between the ETF price and its NAV are associated with lower ETF liquidity as well as a larger difference in the relative liquidity between the ETF and the constituent portfolio (i.e. larger liquidity mismatch). These results support the work of Roll, Schwartz and Subrahmanyam (2007), who argue that market liquidity plays a key role in the price fluctuations that eliminate arbitrage opportunities. We also document a large and significantly positive effect of ETF liquidity on the convergence of the ETF price to its NAV as well as a large and significantly negative effect of liquidity on the duration for which the ETF remains at a premium (positive sign) or at a discount (negative sign) relative to NAV. The effect of liquidity on the ETF premium/discount is stronger when the holding costs of the ETF shares and the underlying portfolio stocks are higher.[2] Our results are robust when we control for the effect of macroeconomic conditions as well as to alternative model specifications.

Whether liquidity has an effect on the size of the ETF premium is difficult to test due to the simultaneity between liquidity and the ETF premium. Our main conjecture is that liquidity facilitates trading. In the same way as Broman and Shum (2018), we show that positive changes in liquidity predict increases in net flows. The reverse causality story – that greater net fund flows lead to higher relative liquidity – while theoretically plausible, is not supported by our results. First, we exclude the first six months after inception for the ETFs in our sample to mitigate issues related to the ETF's cycle. The average age of the ETFs is 5.87 years, indicating that for the majority of funds, their liquidity is already well established. Second, we show that net fund flows do not predict unexpected changes in liquidity.

Our paper is related to a strand of the literature that compares ETFs to open-end mutual funds. Guedj and Huang (2009) argue that open-end fund structures can be viewed as providing insurance for investors exposed to liquidity shocks, and hence is beneficial for risk averse investors. Such liquidity insurance, however, is not without cost as it can create moral hazard issues and reduce the fund performance. They show that open-end funds and ETFs coexist in equilibrium with different liquidity clienteles where ETFs are better suited for narrower and less liquid underlying indexes, and for investors with longer investment horizons.

Finally, we contribute to the literature on limits to arbitrage in international

[2]Our results support the limits to arbitrage hypothesis. For stocks with high holding costs liquidity is important as it allows traders to open and close positions with lower transaction costs. Examples of studies on short sale constraints as limits to arbitrage include Pontiff (1996), Chen, Hong and Stein (2002), Ofek, Richardson and Whitelaw (2004), Chang, Cheng and Yu (2007), or Gromb and Vayanos (2010).

equity markets. Gagnon and Karolyi (2010) examine whether the variation in the magnitude of the deviation from price parity for cross-listed stocks is related to arbitrage costs. Their findings suggest that the deviation is positively related to holding costs, especially idiosyncratic risk, which impedes arbitrage activities. While their study focuses on cross-listed stock pairs, our study identifies the determinants of the persistence and the duration of deviations of ETF prices from their NAVs.

2 Background on ETFs

Exchange-traded funds have grown substantially in size, diversity and market significance in recent years, resulting in increased attention by investors, regulators and academics.[3] This section describes the institutional details of ETF arbitrage and the role of authorized participants (APs). We also discuss why liquidity is an important factor that affects the price difference between the ETF that is traded on a U.S. exchange and the underlying portfolio of international stocks traded in less liquid and more opaque markets.

2.1 ETF Arbitrage Mechanism

ETFs, like closed-end funds, are traded on the secondary market.[4] Inflows to the fund are also not used directly to create new shares. However, unlike the standard closed-end funds, new shares can be created (or redeemed). Institutional investors (called Authorized Participants (APs)), who have entered into a contract with the

[3] In 2016, the assets managed by ETFs globally amounted to approximately USD 3.42 trillion. (Source: ETFGI database).

[4] An ETF is an investment fund regulated under the provisions of the Investment Company Act of 1940 (ICA). As the fund's structure is different from that of a mutual fund, the sponsor must obtain exemption from certain provisions of the Act. These exemptions include the ability to redeem shares only in large blocks, the use of in-kind transactions for creation, and redemption units; and the means for other investment companies to purchase shares of the fund in excess of the ICA's fund-of-fund regulations (Abner, 2010).

ETF issuer, can create a basket of shares (creation unit) replicating the index that the fund tracks, and deliver them to the fund in exchange for ETF shares.[5] In the case of international equity ETFs, a creation unit consists of anywhere from 25,000 ETF shares to 600,000 ETF shares. Alternatively, an AP can redeem ETFs by exchanging one full creation unit of ETF shares for a basket of the underlying securities. This mechanism allows institutional investors to take advantage of arbitrage opportunities when the price of the ETF deviates from its NAV. Therefore, unlike closed-end funds, the mechanism for creating and redeeming ETFs ensures that the premium or discount between the ETF price and the NAV remains in a range equal to the expectation of the current fair value of the portfolio holdings and the associated trading costs.[6]

ETF arbitrage, however, is not riskless and as the risks of arbitrage increase, APs may withdraw from ETF arbitrage activities that correct mispricing and cease to provide liquidity in the ETF market as they are not contractually bound by any legal obligation or monetary incentives from the ETF issuer to create or redeem ETF shares. The next subsection discusses the effect of liquidity on the risk of ETF arbitrage.

[5] APs are the only market participants active in this primary ETF market by institutional design. ETFs typically have an average of 25-35 APs, many of which are also large, leveraged broker-dealer institutions. In contrast to APs, non-AP market participants may trade the ETF intraday on the secondary market but lack the ability to create and redeem shares with the ETF issuers.

[6] The authorized participant must pay a fixed dollar fee ($500 to $28,000), for each creation transaction regardless of the number of creation units involved. For example, for EEM (iShares MSCI Emerging Markets ETF) the fixed fee of $7,700 would currently amount to about 3.7 bp for a single creation unit worth about $21 million. These transaction costs, combined with the costs of trading the underlying securities, would therefore be expected to set the boundaries of how much the ETF price can diverge from its NAV.

2.2 The Effect of Liquidity on ETF Arbitrage

Given the large minimum size of ETF creation/redemption units, APs must be able to efficiently source a variety of international stocks in large quantity via secondary market trades. Given the nature of many international equity markets and the fact that many APs are large financial institutions, APs rely heavily on existing market liquidity as well as their own stock inventory to perform ETF arbitrage. Pan and Zeng (2017), however, argue that this strategy involves an important risk. Since ETF arbitrage is performed in-kind, it may result either in a large price impact or alternatively large shocks to the inventory positions of APs. Arbitrage opportunities are positively correlated with market volatility (increasing the fundamental risk of inventory positions), but negatively correlated with stock market liquidity (the ability to trade arbitraged stocks). Overall, expected liquidity or inventory shocks, which cause capital constraints to bind, can increase the expected returns from arbitrage required by APs.

Kozhan and Tham (2010) also examine the link between the size of the arbitrage deviation and market illiquidity. They find a statistically significant relation in which the deviation is positively correlated with measures of market illiquidity. However, they argue that the economic reason behind this relation is the presence of execution risk as the number of competing arbitrageurs increases. They also study the cost of holding inventory through the profit and loss distribution of arbitrage strategy failures. They find a statistically significant and positive relation between arbitrage deviation and the cost of inventory.

The role of competing arbitrageurs and market illiquidity on the existence of convergence of arbitrage opportunities has been studied recently by, among others, Stein (2009), Kondor (2009), and Oehmke (2009). Kondor (2009) studies the impact of competition on arbitrage whereas Oehmke (2009) investigates the role of strategic arbitrageurs, trading with liquidity frictions, on the speed of price convergence. These papers focus on convergence trading and the use of wealth constraints that cause arbitrageurs to unwillingly unwind their positions as prices diverge. Kozhan and Tham (2010), on the other hand, argue that the underlying reason is the effect of crowded trades on the elimination of arbitrage.

3 Data and Summary Statistics

3.1 Sample Selection

We use different datasets in generating our sample. We start by searching the ETF Global database for all non-U.S. equity ETFs. We focus on ETFs that are traded in the U.S. but replicate international equity indices. We exclude leveraged, inverse and synthetic ETFs where the fund issuer enters into a total return swap with a financial institution that promises to deliver the performance of the index to the fund (see Ramaswamy, 2011).

For each ETF, we obtain daily data on its constituents' identifiers and portfolio weights. We remove all fund-day observations for which the weights of constituent stocks are missing or the total weight exceeds 100%. We exclude all ETFs that have funds allocated to assets traded over-the-counter (such as bonds), to shares in privately held companies or to derivative securities such as warrants, rights or stock options. We treat foreign currency contracts as cash and assign zero illiquidity values to them.

We obtain trading data for the ETFs from CRSP. We remove observations for which the bid-ask spread is negative, as well as observations during the first 180 days since the inception date of the ETF. We also collect daily price and volume data for each constituent stock from Datastream. Our final sample consists of 584 International Equity ETFs for the period January 3, 2012 to December 29, 2017.

We begin the study by computing the premium/discount of the ETF relative to its NAV. The ETF premium or discount for fund f at time t is calculated as the ratio of price-NAV difference over its NAV:

$$premium_{f,t} = \frac{P_{f,t} - NAV_{f,t}}{NAV_{f,t}}$$

3.2 Measures of Illiquidity

Most of the empirical illiquidity measures require TAQ-type data, which is not available for foreign markets. The Amihud (2002) measure of price impact has the advantage that it only requires daily data on trading volume and asset price, which are readily available even for emerging markets. In addition, Hasbrouck (2003) finds the Amihud measure to be highly correlated with the TAQ-based price impact measure in the U.S. market.

First, we calculate a monthly measure of illiquidity (as of the last trading day of each month), for each individual stock calculated as the average ratio of the daily absolute return to the daily trading volume within that month.[7] The Amihud illiquidity measure for stock i in month t for ETF f, $IL_{i,f,t}$, is defined as:

$$IL_{i,f,t} = \frac{1}{D_t} \sum_{d=1}^{D_t} |r_{i,d}|/VOL_{i,d}$$

where D_t is the number of trading days in month t (approximately 21 days), $r_{i,d}$ and $VOL_{i,d}$ are stock i's daily return and volume in day d of month t.[8] The ratio $|r_{i,d}|/VOL_{i,d}$ is the absolute proportional price change per unit of daily trading volume, or equivalently the daily price impact of the order flow. This is closely related to Kyle's (1985) concept of illiquidity defined as the response of price to order flow.

Second, the individual stock measure of illiquidity is then aggregated to compute a measure of the ETF's underlying portfolio's monthly illiquidity $AIL_{f,t}$ as of the last trading day of each month:

$$AIL_{f,t} = \sum_{i=1}^{N_{f,t}} w_{i,f,t} \times IL_{i,f,t}$$

[7] We construct the illiquidity measure using dollar absolute returns and dollar volumes. We obtain similar results if we use local currency prices and daily volume.

[8] Unlike Amihud (2002) where illiquidity is calculated annually with at least 200 daily data each year, we only use 21 days to calculate monthly value for the measure. This is because we would like to use non-overlapping data in constructing our illiquidity measure so that we can relate this measure to ETF premiums/discounts at a monthly frequency.

where $N_{f,t}$ is the number of stocks included in fund f in month t and $w_{i,f,t}$ is the weight of stock i included in fund f in month t. Finally, we construct the Amihud illiquidity measure for each ETF fund f, $FIL_{f,t}$, since each ETF itself is a traded instrument in the U.S. market.

3.3 Summary statistics

Table 1 presents the cross-sectional characteristics for the sample of international equity ETFs used in this study. The median fund's age is five years and its market capitalization is $62 million in assets, but the distribution is heavily skewed in terms of size with the largest fund (iShares MSCI EAFE) accounting for around $80 billion. Dollar trading volume is even more skewed, with the median fund trading around $20,000 per day and the most active fund (iShares MSCI Emerging Markets ETF) trading around $21 million per day. The median ETF bid-ask spread is just over 18 basis points (bp), but it varies from as low as 1 bp for the most liquid funds to as much as several percent for the least liquid funds, reflecting the wide disparity in trading costs across funds. The annual expense ratio of the median fund amounts to 56 bp of net assets, varying from 4 bp to 643 bp across funds. The median fund's creation unit is 50,000 shares and the creation fee is $1,700.

Panel A of the table also shows the distribution of ETF characteristics by development level. The emerging markets funds are smaller, more illiquid (larger bid-ask spreads and lower trading volume) and have higher expense ratios. Panel B shows the distributions of the same characteristics by investment category. Broad equity-based funds are the largest, the most liquid and the cheapest in terms of expense ratio. These funds, however, have a higher creation fee.

Table 2 reports the same cross-sectional characteristics as in Table 1 for sub-samples of funds split by region and country. Panel A of the table shows that the majority of ETFs in our sample have underlying assets traded in Europe and Asia-Pacific. Latin America and Middle-East Africa ETFs account for less than a third and a tenth respectively of the number of Europe/Asia-Pacific ETFs. The median fund's age ranges from five years (Europe ETFs) to nine years (Middle-East Africa ETFs). As expected, daily volume and turnover are the highest and trading costs are the lowest for Europe ETFs with a median bid-ask spread of 16 bp versus 55 bp for the other four regions. Compared to U.S. equity funds, Europe ETFs display a smaller median size and smaller daily volume and turnover. The bid-ask spread and expenses, however, are very close (see Petajisto, 2017). Median fund age is close to the one reported by Hilliard (2014).

Panel B of Table 2 shows the same characteristics for the four countries with the largest number of ETFs in our sample. 74% of ETFs in our sample track multi-country indices with half of these funds tracking global markets and a fifth emerging markets. 26% of the sample consists of country ETFs with China accounting for 20% or 33,260 fund-year observations. The median fund's trading volume and turnover are the largest for India ETFs and the lowest for Japan and China ETFs.

Table 3 shows statistics on share creations and redemptions for the sample period. We compute the fraction of trading days when each ETF experienced share creation/redemption, and then compute the mean across all funds. On average, creations occurred on 2.02% of days whereas redemptions occurred only on 1.23% of all trading days. However, these measures are heavily skewed by many small funds with little or no activity.[9] The other columns in the table are all conditional on creations or redemptions taking place on a given day. The median number of ETF shares created (redeemed) was 250,000 (200,000) which is five (four) times the median creation unit. The table, however, shows that the size of the creations and redemptions varied with fund development level and category. Developed markets ETFs (development level) and Broad Equity ETF (category) have the largest number of shares created and redeemed as well as the largest creation and redemption values. The median dollar value of these transactions for all funds was \$8.57 million for creations and \$6.34 million for redemptions whereas for Developed markets funds they were \$11.53

[9]Panel C of Table A1 in the appendix shows that U.S. equity funds on average experienced non-zero net flows around 20% of trading days during our sample period.

million and $7.94 million respectively and for Broad Equity, the values were $13.02 million and $9.66 million. In comparison the median U.S. equity ETFs created or redeemed 100,000 shares for a median value of $ 4 million and a mean transaction size of 4,727% of the daily trading volume (see Petajisto, 2017).

As a fraction of the daily ETF trading volume, the median creation or redemption transaction accounted for more than 12,500% of daily volume. Economically, these numbers indicate that the size of a creation unit is very large for a typical international equity ETF. Even if an arbitrageur participates in every single trade in a fund and always on the same side, in most funds he would still take several days for her to accumulate a position that would be large enough to offset the creation or redemption of a single creation unit. This introduces timing and execution risk which makes it harder to arbitrage small mispricings by using the ETF share creation and redemption process.

Table 4 shows the number of ETFs in each investment category. Our sample covers 584 funds (including active funds and such that underwent closure) over the sample period, with a total of $4.7 billion in assets. The largest categories are broad equity funds with $1.3 billion in assets.[10] The statistics on the ETF premiums are computed following Petajisto (2017). First, we calculate the average level and time-series volatility of the premium for each fund. Then, we average across funds within each category to calculate the average premium and the average volatility of the premium. To compute value-weighted averages, we weight each fund by its mean market capitalization over the sample period.

The average premium is only 9 bp, however, the time-series volatility of that premium is 36 bp, which indicates that ETF prices fluctuate considerably around NAVs. The value-weighted average volatility is comparable at 35 bp, so the result is not limited to smaller funds. Economically, the equal-weighted number tells us that the typical fund is trading in a range from -61.56 bp to 79.56 bp around its

[10]The original sample contained one growth and one value ETFs (two of Global X EM funds) but observations were available only for the first 30 days in 2012. After that the funds were withdrawn from the market.

NAV with a 95% probability. Petajisto (2017) and Hilliard (2014) report an average premium of 0 bp and a volatility of 18 bp for U.S. funds.

To give some idea of the transaction costs in the ETFs themselves, the last two columns of Table 4 show the bid-ask spreads for the funds in our sample. The equal-weighted average is 48 bp while the value-weighted average is 44 bp. The bottom two panels of the table show the same statistics across funds sorted into quintiles by trading volume and Amihud's illiquidity measure. As expected, the ETFs with the highest volume and the lowest price impact (illiquidity measure) have the largest market cap, the lowest ETF premium, and the smallest bid-ask spread.

Finally, Table 5 reports descriptive statistics on ETF premiums and trading costs by region and country of the underlying assets of the funds. Panel A shows that the median market capitalization is $29 million, however, the distribution is skewed with a large average fund size and standard deviation. The largest funds are the EAFE (Europe, Australasia and Far East) ETFs as well as emerging and developed markets ETFs. Median bid-ask spread is at 36 bp but ranges from as low as 11 bp for EAFE funds to as high as 75 bp for Middle-East Africa funds. The average ETF premium is 11 bp with a time-series volatility of 36 bp. There is large variation across regions with a 30 bp discount for Middle-East Africa ETFs and a 21 bp premium for EAFE ETFs.

Panel B shows that country ETFs are on average smaller and have a lower ETF premium and bid-ask spread than multi-country ETFs. There is, however, substantial variability. There are several very large Japan and Germany ETFs skewing the average market capitalization. The average premium is 6 bp, however, five countries display a median discount (as low as -16 bp for Indonesia), while all other country ETFs trade at a premium relative to their NAV (up to 38 bp for Russia ETFs). Economically, the premium of a country ETF is between -86.02 bp and 97.84 bp with 95% probability. The next section presents the results from our formal analysis.

4 Empirical Analysis

4.1 ETF Premium and Illiquidity: Baseline model

In this section, we test how variation in the fund premium/discount is related to illiquidity. Our main conjecture is that liquidity facilitates trading. In this case, liquidity not only should be an important determinant of the size of the ETF premium or discount, but liquidity should also have a positive impact on the fund flows. In addition, as argued by Pan and Zeng (2019), a greater liquidity mismatch (higher relative liquidity) between the ETF and its underlying portfolio should result in higher and more persistent mispricing. On the other hand, when both the ETF and the underlying assets are liquid, this will facilitate trading and therefore eliminate or reduce mispricing.[11]

Our first regression specification is:

$$|premium|_{f,t} = \alpha_f + \beta_f \times Illiquidity_{f,t} + \gamma_f \times Fund\ control\ variables_{f,t} + \epsilon_{f,t} \quad (1)$$

where $|premium|_{f,t}$ is the absolute value of the ETF market price at time t over its NAV minus one; $Illiquidity_{f,t}$ is the ETF illiquidity ($FIL_{f,t}$), the illiquidity of the underlying portfolio ($AIL_{f,t}$) or the relative (ETF over the underlying portfolio) illiquidity measure. We use monthly averages of all the variables. Although, the arbitrage activities of APs often occur at the daily frequency, Broman and Shum (2018) find that while the daily premiums only persist for a few days, the average premiums over a monthly horizon are persistent. Given that the underlying assets in our sample are traded on (less liquid and efficient) non-U.S. markets, it is likely that the ETF premium/discount for our sample of ETFs will be even more persistent.[12]

[11] We focus on both the ETF liquidity as well as on the liquidity of the underlying portfolio of securities since a successful arbitrage strategy requires that an institutional investor have the ability to trade easily both the ETF and its underlying portfolio.

[12] Panel C of Table A1 shows that the percentage of days with non-zero net flows to international equity ETFs is less than half the percentage for U.S. equity ETFs.

The purpose of our paper is to explain the size of this persistent component of the premium.[13]

$Fund\ control\ variables_{f,t}$ is a vector of ETF-specific characteristics, including size (the natural logarithm of a fund's assets under management), age, expense ratio and listing exchange. We use fund-year fixed effects in all specifications. Controlling for fund fixed effects is particularly important in order to ensure that our results are not driven by omitted fund characteristics that might be correlated with the funds' liquidity characteristics. We also control for Pastor and Stambaugh (2003) aggregate level of liquidity as well as the TED spread and VIX index as proxies for the availability and cost of arbitrage capital. The variable definitions are in Table A2 in the appendix to this paper.

Table 6 reports the results from the estimation of equation (1).[14] The coefficient of the ETF illiquidity measure has the expected sign and remains statistically significant when we control for ETF characteristics, limits to arbitrage, and the macroeconomic conditions. Larger ETF Amihud illiquidity values are associated with larger absolute values of the ETF premium/discount. For example in column (2), a one standard-deviation increase in the ETF illiquidity measure results in an increase in the absolute value of the ETF premium of 0.55%. This is statistically significant and economically large given that the mean (median) premium is seven (nine) basis points. Table 6 provides no support that the underlying portfolio illiquidity has a significant effect on the absolute value of the ETF premium. The coefficients are not statistically significant, and the magnitude of the effect is much smaller than the effect of the ETF illiquidity.[15]. From columns (3) and (4), however, the decrease in the liquidity

[13]We calculate the half-life of the deviations from NAV as a measure of persistence. We define an autoregressive process of order one for the log of the ratio between the daily ETF price and its NAV and calculate the horizon at which the percentage deviation from parity is half. The autoregressive coefficient is 0.89 which implies a half-life of 5.95 trading days.

[14]We estimate equation (1) both in levels and in differences in order to account for the persistence in the premium and the illiquidity measures. The results remain the same.

[15]The weaker results for underlying portfolio illiquidity is potentially due to the lack of variations because many ETFs in our sample track similar underlying portfolios

mismatch between the ETF and the underlying portfolio (the smaller the gap between FIL and AIL) has a large and significant negative effect on the absolute value of the ETF premium. This finding is in line with Pan and Zeng (2019) who suggest that for assets with significant liquidity mismatch there could be persistent mispricing. The coefficients of the control variables have the expected sign. For example, the table shows that larger, older funds and funds with a higher expense ratio tend to be priced more efficiently. However, only fund size and expense are significant.

Table 7 reports the results from the estimation of equation (1) for sub-samples of ETFs by market development (Developed vs Emerging markets ETFs). The coefficient of the ETF illiquidity has the expected sign for both sub-samples but it is significant only for the subsample of emerging markets ETFs. The magnitude of the liquidity effect also differs by market development as liquidity has a stronger impact on the ETF premium for emerging markets ETFs. The differences between the illiquidity coefficient (FIL) and the liquidity mismatch coefficient (FIL/AIL) for emerging versus developed markets ETFs are significant at conventional levels (see the coefficients of the interaction variables in Table 7). These results provide further support for the importance of liquidity for implementing arbitrage trading when assets are thinly traded.[16]

4.2 Fund Flows and Liquidity

In this subsection, we address the potential endogeneity problem arising between illiquidity and the ETF premium or discount. Liquid assets may attract more arbitrage trades that reduce deviations from the NAV, while arbitrage opportunities (large deviations from NAV) may themselves attract more liquidity to the market

[16] We estimate specification (1) for subsamples of ETFs split by region (Europe, Asia-Pacific, Latin America and Middle-East and Africa). The coefficients of the illiquidity measures remain similar to the coefficients in Table 6. The coefficients of the interaction variables between illiquidity and region, however, are not significant at conventional levels. The results are available upon request.

(reverse causality). To test for the linkage between illiquidity and NAV deviations we hypothesize that liquidity facilitates trading. In that case, we expect that changes in liquidity will predict net fund flows. Under the alternative hypothesis (the reverse causality), however, net fund flows will predict unexpected changes in liquidity.

We define monthly net fund flows, $NFL_{f,t}$ as:

$$NFL_{f,t} = \frac{\sum_{i=1}^{D_t}(SHR_{f,d} - SHR_{f,d-1}) \times NAV_{f,d}}{AUM_{f,t-1}}$$

where D_t is the number of trading days in month t, $SHR_{f,d}$ the number of shares outstanding for ETF f on day d and $AUM_{f,t-1}$ assets under management for ETF f in month $t-1$.[17]

We estimate the following specification:

$$NFL_{f,t} = a_f + b_f \times Illiquidity_{f,t} + c_f \times Fund\ control\ variables_{f,t} + e_{f,t} \quad (2)$$

where $NFL_{f,t}$ is the net fund flow at time t for fund f; $Illiquidity_{f,t}$ is the ETF illiquidity ($FIL_{f,t}$), the illiquidity of the underlying portfolio ($AIL_{f,t}$) or the relative (ETF over the underlying portfolio) illiquidity measure. We include the past three months' return to the vector of control variables to capture the effect of past performance on fund flow.

The results from specification (2) are reported in Table 8 columns (1) and (2), and they suggest that ETF illiquidity is indeed negatively and significantly related to net fund flows, while the underlying portfolio illiquidity does not have a significant effect.[18] This result cannot be explained by the reverse causality hypothesis. Our

[17] Broman and Shum (2018) argue that this definition is more accurate because ETF flows are measured using actual net asset values observed on the day when flows occurs. In contrast, the monthly flow variable typically used in the mutual fund literature (i.e. $(AUM_{f,t} - AUM_{f,t-1}) \times (1 + r_{f,t})$), assumes that flows occur at the end of each period (see for example, Sirri and Tufano, 1998). The difference is due to the fact that mutual fund data are available only at the monthly interval, unlike ETF data which can be obtained daily.

[18] Although, the adjusted R^2 in Table 8 are small, they are in line with prior studies on the relative illiquidity of ETFs and their underlying constituents (see Clifford, Fulkerson and Jordan, 2014 or Broman and Shum, 2018)

findings are in line with evidence on U.S. equity ETFs. Broman and Shum (2018) document a significantly positive association between ETF liquidity and net fund flows. We also investigate the predictive power of unexpected changes in illiquidity on future fund flows. To measure unexpected changes in liquidity, we first extract the residuals from an AR(1) model estimated separately for each illiquidity variable and the fund characteristics in equation (1). The results in Table 8 column (3) and (4) show that shocks to ETF illiquidity continue to predict net fund flows negatively at conventional levels of significance.

Given the rapid growth and relatively young age of the ETF industry, the flow-liquidity relation may vary over time. We carry out cumulative sum (of recursive residuals) tests for parameter stability. The null hypothesis of no structural break cannot be rejected at conventional confidence levels.

Next, we examine the ability of net fund flows to predict future changes in illiquidity. We include the same control variables as in the fund-flow equation (2). Our results suggest that net fund flows is insignificant in predicting future changes in liquidity.[19] Thus, investor demand does not seem to have significant predictive power over future changes to liquidity in our sample of funds.

Our results remain robust to alternative specifications. Table A3, in the appendix, reports the results from a bivariate panel VAR model of net fund flows and ETF illiquidity.[20] The results are consistent with the results in Table 8 where positive changes in liquidity predict increases in net flows. Net fund flows, however, do not predict unexpected changes in liquidity. These results are supported by bivariate Granger causality tests suggesting that causality runs from illiquidity to net fund flows.[21]

[19] The results are available upon request.

[20] Estimating the reduced form panel VAR model does not require any further assumptions. Calculating impulse response functions, on the other hand, requires recovering of the structural parameters.

[21] We thank the referee for suggesting the panel VAR approach to evaluating the relation between illiquidity and net fund flows.

4.3 Duration analysis

In this subsection, we carry out a duration analysis and examine how liquidity affects the length of time during which (1) an ETF's market price and its NAV converge and (2) an ETF price remains on the same side of its NAV. We use a standard Cox proportional hazard regression framework to estimate the coefficients and the effect of illiquidity on the conditional probability that a "failure event" occurs. We define the "failure event" as (1) the convergence of the ETF price and its NAV and (2) the change in sign of the ETF premium/discount (the ETF price moving to the opposite side of its NAV).

First, we calculate the differential between the two prices as outlined below and define the convergence of the two prices as when the price differential is smaller than the estimated round-trip trading costs.

$$price\ diff_{f,t} = \frac{|P^{ETF}_{f,t} - NAV_{f,t}|}{(P^{ETF}_{f,t} + NAV_{f,t})/2} \tag{3}$$

When the price differential is small relative to the trading costs, it may not be worthwhile for investors to trade to take advantage of the deviation from NAV. For a long-short strategy, there are two times the roundtrip transaction costs as positions are opened and closed on both the long and the short side. Investors taking one-sided positions incur at least one round-trip transaction cost. Grundy and Martin (2001) calculate the raw and risk-adjusted returns of a zero-investment momentum trading strategy and estimate that a 1.5% round-trip cost would make the profits insignificant. Mitchell, Pulvino and Stafford (2002) assess the effect of transaction costs on the returns of risk arbitrage portfolios. By comparing the returns of a value-weighted-average-return portfolio and a risk-arbitrage-index-manager portfolio, they estimate an approximately 1.5% reduction in annual return by direct transaction costs (commission, surcharges, taxes) and another 1.5% reduction by indirect transaction costs (price impact). In sharp contrast, Frazzini, Israel and Moskowitz (2012) use a unique dataset of live trades from a large institutional investor to examine the trading costs, net of cost returns, and break even fund sizes of equity strategies and they find that the trading cost estimates are many times smaller and the fund sizes are more than an order of magnitude larger than those reported in previous studies.[22]

Frazzini's et al. (2012) mean value-weight price impact and implementation shortfall cost for international portfolios is 20 basis points, which we use as a round-trip trading cost. This defines the first "failure event" as the $price\ diff$ is smaller than 20 bp , i.e. the convergence of an ETF market price to its NAV. The second failure event is the change in the sign of the ETF premium.

We estimate the following Cox proportional hazard model:

$$h(t) = h_0(t)e^{(A_{f,t})} \tag{4}$$

where $h(t)$ is the hazard ratio, $h_0(t)$ is the baseline hazard, and $A_{f,t}$ is an array of explanatory variables including ETF and the underlying portfolio illiquidity and the fund control variables in eq(1):

$$A_{f,t} = \gamma_1 \times Illiquidity_{f,t} + \gamma_2 \times Fund\ Control\ Variables_{f,t} \tag{5}$$

Table 9 presents the estimation results from a duration model for the ETF price convergence to its NAV. The ETF illiquidity is negatively associated with the hazard ratio. The coefficients are negative, statistically significant and economically large. A one standard-deviation decrease in the ETF illiquidity is associated with an average of 1.34% increase in the conditional probability of the price convergence. On the other hand, the underlying portfolio liquidity has no strong effect on the hazard ratio. These results are consistent with those in Table 6.

Next, Table 10 presents the estimation results from a duration model for the number of trading days that the ETF trades only at a premium or only at a discount relative to its NAV. The average length for which the ETF premium has the same sign is 5.4 trading days. The median is two trading days, however, for almost 15% of the observations, the ETF premium has the same sign for longer than 10 trading

[22]They use actual trading costs from a real-world arbitrageur to estimate price impact rather than aggregated trade and quote level data used in other studies.

days (two calendar weeks). We require a measure of illiquidity at daily frequency so we use the ETF bid-ask spread and the bid-ask spread of the portfolio of constituent stocks. The coefficients of ETF illiquidity are negative, statistically significant and economically large. Similar to the results in Table 9, the underlying portfolio bid-ask spread does not have a significant effect on the hazard ratio. The rest of the explanatory and control variables have the expected sign.

Finally, we explore a possible channel through which liquidity may affect the convergence of the ETF market price to its NAV. Liquidity may increase the probability of price convergence by attracting arbitrage trading as stock liquidity lowers the transaction costs and eases arbitrage trading. With this channel, the effect of liquidity should be stronger for ETFs with high holding costs, because liquidity allows traders to trade at lower transaction costs and is particularly important for these ETFs. Using idiosyncratic volatility as a proxy for holding cost, we test this channel by augmenting equation (5) with the interaction terms between the ETF illiquidity and the idiosyncratic volatility. The liquidity effect on the price convergence may also be stronger where the idiosyncratic volatility is high, because noise trading and speculative trading provide liquidity and increase the idiosyncratic volatility (Harris, 1989, Greene and Smart, 1999, and Foucault, Sraer and Thesmar, 2011).

Our results suggest that the ETF illiquidity is negatively associated with the conditional probability of the convergence. The coefficients for the interaction terms between the ETF illiquidity and its idiosyncratic volatility have the same sign as the coefficients for the illiquidity variable. One standard-deviation decrease in ETF illiquidity corresponds to a 3.62% increase in the hazard ratio for an ETF with the mean idiosyncratic volatility, and a 4.16% increase in the hazard ratio for an ETF with an idiosyncratic volatility that is one standard-deviation above the mean. Overall, the results indicate that the liquidity effect on the conditional probability of the price convergence is stronger where the idiosyncratic volatility is high, which provides some support for the limits to arbitrage hypothesis, i.e. that liquidity increases the conditional probability of the convergence because of high holding costs.

5 Conclusions and discussions

The market for ETFs, and especially international ETFs, has grown into a $5 trillion market over the last decade. Its popularity stems from investors searching for an instant and low cost access to a diversified portfolio of stocks and investment styles in an international setting. The stock-like features of ETFs, with their real-time trading and high liquidity (accounting for one quarter of the daily volume on U.S. stock markets) are an additional advantage compared to mutual funds. The creation and redemption of ETF shares by Authorized Participants adds to this liquidity and allows for an efficient arbitrage mechanism ensuring that ETF prices remain very close to their NAV. Evidence on this phenomenon has been documented for U.S. equity ETFs (see Elton, Gruber, Comer and Li (2002) and Engle and Sarkar (2006)). This may, however, not be true for international equity ETFs due to limits to arbitrage which hinder the creation or redemption of ETF shares by APs.

This paper examines the relationship between an ETF premium or discount and market liquidity. Using a sample of international equity ETFs that trade on U.S. stock exchanges, we examine the determinants of the cross-sectional variation in the (absolute) size of the premium or discount, and the price convergence to NAV. Our results show that lower ETF liquidity leads to a larger price difference between the ETF price and the value of the underlying securities. The effect remains significant after we control for fund characteristics, holding costs and the overall macroeconomic conditions. This relation is particularly important for thinly traded ETFs such as the Emerging market funds in our sample.

We also estimate two types of duration models. First, we examine the effect of liquidity on the length of time during which the difference between the ETF price and the NAV of its underlying securities exceeds the trading costs of a round-trip arbitrage strategy. Our results show that there is a positive relationship between ETF liquidity and the conditional probability of the price convergence.

Second, we estimate a duration model for the number of trading days that the ETF trades only at a premium or only at a discount relative to its NAV. The average length for which the ETF premium has the same sign is 5.4 trading days. The median is two trading days, however, for almost 15% of the observations, the ETF premium has the same sign for longer than 10 trading days (two calendar weeks). We require a measure of illiquidity at daily frequency so we use the ETF bid-ask spread and the bid-ask spread of the portfolio of constituent stocks. The coefficients of ETF illiquidity are negative, statistically significant and economically large. Finally, we find that idiosyncratic volatility strengthens the liquidity effect. Overall our results show that liquidity is an important determinant of the size of the ETF premium. They also shed light on the way in which deviations from its NAV are eliminated.

An important implication of our results is that international ETFs constitute an efficient way for U.S. investors to gain exposure to foreign equity markets. The high liquidity of ETFs and their arbitrage mechanism ensure that ETFs better reflect prices of the underlying international stocks. This is especially true compared to international closed-end funds, which on average trade at a discount of 7.9% as reported by Cherkes, Sagi and Stanton (2009). Important research questions remain to be answered. For example, given that premiums for ETFs relative to their NAVs are very different for comparable ETFs (i.e. ETFs with similar baskets of underlying assets) and there is only a gradual converge of prices to NAV values, is there a potential wealth transfer between uninformed retail and informed institutional investors.

References

[1] Abner, D. (2010). The ETF handbook: how to value and trade exchange traded funds. *Wiley Finance*, 2nd Edition

[2] Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets*, 5(1), 31-56.

[3] Ben-David I., F. Franzoni and R. Moussawi (2018). Do ETFs increase volatility?, *The Journal of Finance*, 73(8), 2471-2535.

[4] Bhattacharya, A. and M. O'Hara (2018). Can ETFs increase market fragility? Effect of information linkages in ETF markets, *Working paper.*

[5] Broman M. and P. Shum (2018). Relative liquidity, fund flows and short-term demand: Evidence from exchange-traded funds, *Financial Review*, 53(1), 87-115.

[6] Chacko, G., S. Das and R. Fan (2016). An index-based measure of liquidity, *Journal of Banking and Finance*, 68, 162-178.

[7] Chang, E., J. Cheng and Y. Yu (2007). Short-sales constraints and price discovery: Evidence from the Hong Kong market, *The Journal of Finance*, 62(5), 2097-2121.

[8] Chen, J., H. Hong and J. Stein (2002). Breadth of ownership and stock returns, *Journal of Financial Economics*, 66(2/3), 171-205.

[9] Cherkes, M., J. Sagi and R. Stanton (2009). A liquidity-based theory of closed-end funds, *The Review of FInancial Studies*, 22(1), 257-297.

[10] Clifford, C., J. Fulkerson and B. Jordan (2014). What drives ETF Flows?, *Financial Review*, 49(3), 619-642.

[11] Cong, W. and D. Xu (2016). Rise of factor investing: asset prices, informational efficiency, and security design, *Working paper.*

[12] Dannhauser, C. (2017). The Impact of Innovation: Evidence from corporate bond ETFs, *Journal of Financial Economics*, 125(3), 537-560.

[13] Duan, Y., G. Hu and D. McLean (2010). Costly arbitrage and idiosyncratic risk: Evidence from short sellers, *Journal of Financial Intermediation*, 19(4), 564-579.

[14] Elton, E.J., M.J. Gruber, G. Comer and K. Li (2002). Spiders: Where are the bugs?, *Journal of Business*, 75(3), 453-472.

[15] Engle, R. and D. Sarkar (2006). Premiums-discounts and exchange traded funds, *Journal of Derivatives*, Summer, 27-45.

[16] Foucault, T., D. Sraer and D. Thesmar (2011). Individual investors and volatility, *The Journal of Finance*, 66(4), 1369-1406.

[17] Frazzini, A., R. Israel and T. Moskowitz (2012). Trading costs of asset pricing anomalies, *Working paper.*

[18] Gagnon, L. and A. Karolyi (2010). Multi-market trading and arbitrage, *Journal of Financial Economics*, 97(1), 53-80.

[19] Gastineau, G. (2010). The exchange-traded funds manual, *John Wiley and Sons*, 2nd Edition.

[20] Greene, J. and S. Smart (1999). Liquidity provision and noise trading: Evidence from the "investment dartboard" column, *The Journal of Finance*, 54(5), 1885-1899.

[21] Gromb, D. and D. Vayanos (2010). Limits of arbitrage, *Annual Review of Financial Economics*, 2, 251-275.

[22] Grundy, B. and J. Martin (2001). Understanding the nature of the risks and the source of the rewards to momentum investing, *Review of Financial Studies*, 14(1), 29-78.

[23] Guedj, I., and J. Huang (2009). Are ETFs replacing index mutual funds?, *Working paper.*

[24] Hamm, S. (2010). The effect of ETFs on stock liquidity, *Working paper.*

[25] Harris, L. (1989). S&P 500 cash stock price volatilities. *The Journal of Finance*, 44(5), 1155-1175.

[26] Hasbrouck, J. (2003). Intraday price formation in U.S. equity index markets, *The Journal of Finance*, 58(6), 2375-2400.

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

[28] Kondor, P. (2009). Risk in dynamic arbitrage: Price effects of convergence trading, *The Journal of Finance*, 64(2), 631–655.

[29] Kozhan R. and W. Tham (2010). How riskless is "riskless" arbitrage?, *Working paper.*

[30] Kyle, A. (1985). Continuous auctions and insider trading, *Econometrica*, 53, 1315-1335.

[31] Malamud, S. (2016). A dynamic equilibrium model of ETFs, *Working paper.*

[32] Mitchell, A., T. Pulvino and E. Stafford (2002). Limited arbitrage in equity markets, *The Journal of Finance*, 57(2), 551-584.

[33] Oehmke, M. (2009). Gradual arbitrage, *Working paper.*

[34] Ofek, E., M. Richardson and R. Whitelaw (2004). Limited arbitrage and short sales restrictions. Evidence from the options markets, *Journal of Financial Economics*, 74(2), 305–342.

[35] Pan, K. and Y. Zeng (2019). ETF arbitrage under liquidity mismatch, *Working paper*.

[36] Pastor, L. and R. Stambaugh (2003). Liquidity risk and expected stock returns, *Journal of Political Economy*, 111(3), 642-685.

[37] Petajisto, A. (2017). Inefficiencies in the pricing of exchange-traded funds, *Financial Analysts Journal*, 73(1), 24-54.

[38] Pontiff, J. (1996). Costly arbitrage: Evidence from closed-end funds, *Quarterly Journal of Economics*, 111(4), 1135-1151.

[39] Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk, *Journal of Accounting and Economics*, 42(1/2), 35-52.

[40] Ramaswamy, S. (2011). Market structures and systemic risks of exchange-traded funds, *Working paper*.

[41] Richie, N. and J. Madura (2007). Impact of the QQQ on liquidity and risk of the underlying stocks, *Quarterly Review of Economics and Finance*, 47(3), 411-421.

[42] Roll, R., E. Schwartz, and A. Subrahmanyam (2007). Liquidity and the law of one price: the case of the futures-cash basis, *The Journal of Finance*, 62(5), 2201-2234.

[43] Shleifer, A. and R. Vishny (1997). The limits of arbitrage, *The Journal of Finance*, 52(1), 35-55.

[44] Sirri, E. and P. Tufano (1998). Costly search and mutual fund flows, *The Journal of Finance*, 53(5), 1589-1622.

[45] Stein, J. (2009). Sophisticated investors and market efficiency, *The Journal of Finance*, 64(4), 1517-1548.

[46] Tucker, M. and S. Laipply (2010), Understanding bond ETF premiums and discounts: A conceptual framework, *Investment Insights: The Journal from Blackrock*, 13, 1-20.

Table 1: Cross-sectional Sample Statistics for ETFs

The table presents cross-sectional sample statistics for 584 U.S.-listed International Equity ETFs for the period January 3, 2012 to December, 29 2017. Daily volume and turnover represent the trading by investors in ETF shares, and they are computed as the mean throughout the sample period. The bid-ask spread of an ETF is computed as the mean end-of-day closing spread. Market capitalization is the average of the last available month-end value for each year.

Panel A: Development level	All Funds			Blended Markets			Developed Markets			Emerging Markets		
	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median
Age (years)	5.87	4.45	5	5.64	3.54	5	6.86	5.68	5	5.19	3.66	4
Market cap ($millions)	1,014	4,942	62	656	3,463	66	1,357	5,747	73	1,077	5,427	52
Daily volume ($millions)	0.83	4.95	0.02	0.52	4.88	0.01	0.76	3.67	0.02	1.23	6.52	0.01
Daily turnover (%)	0.07	0.56	0.02	0.06	0.27	0.01	0.07	0.63	0.02	0.08	0.71	0.02
Bid-ask spread (bp)	39.47	55.33	18.89	41.57	56.22	21.54	30.49	42.06	15.12	45.52	60.52	20.99
Expenses (bp)	55	20	58	56	21	56	46	17	48	64	16	64
Creation unit (# of shares)	79,893	74,432	50,000	68,804	40,714	50,000	93,800	97,332	50,000	77,409	71,775	50,000
Creation fees ($)	3,317	3,952	1,700	2,614	3,813	1,000	3,911	4,245	2,000	3,421	3,691	2,000
Observations		627,863			189,997			198,117			179,061	

Panel B: Category	Broad Equity			Sector			Size and Style			Strategy		
	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median
Age (years)	6.41	5.38	5	5.18	3.20	5	6.23	4.15	6	5.16	3.11	5
Market cap ($millions)	1,788	6,800	101	434	3541	37	483	1093	60	256	638	54
Daily volume ($millions)	1.28	6.10	0.02	0.65	5.44	0.01	0.44	2.72	0.01	0.07	0.19	0.01
Daily turnover (%)	0.07	0.63	0.02	0.10	0.68	0.02	0.05	0.21	0.01	0.04	0.21	0.01
Bid-ask spread (bp)	34.21	49.93	14.42	50.19	63.44	25.67	38.28	51.59	19.79	36.39	44.75	22.50
Expenses (bp)	53	20	50	59	21	58	54	18	58	59	14	58
Creation unit (# of shares)	102,722	100,548	50,000	52,996	14,869	50,000	70,792	43,285	50,000	61,771	20,488	50,000
Creation fees ($)	4,490	4,416	2,600	1,261	1,153	1,000	4,439	4,916	2,500	1,837	1,725	1,000
Observations		255,376			153,697			91,131			79,238	

Table 2: Sample Statistics for International Equity ETFs by Region and Country

The table presents statistics on the four most frequent regions and countries in which U.S.-listed International Equity ETFs were invested for the period January 2012 to December 2017. Daily volume and turnover represent the trading by investors in ETF shares, and they are computed as the mean throughout the sample period. The bid-ask spread of an ETF is computed as the mean end-of-day closing spread. Market capitalization is the average of the last available month-end value for each year.

Panel A: Region	Europe			Asia-Pacific			Latin America			Middle-East Africa		
	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median
Age (years)	5.91	4.77	4.90	5.69	3.60	5.02	5.72	2.88	6.01	7.58	1.40	8.97
Market cap ($millions)	1,137	2,228	34	480	879	43	334	609	50	72	2	69
Daily volume ($millions)	0.64	1.46	0.01	0.08	0.17	0.01	0.09	0.18	0.02	0.00	0.00	0.01
Daily turnover (%)	0.27	0.96	0.03	0.09	0.34	0.01	0.04	0.06	0.02	0.00	0.00	0.01
Bid-ask spread (bp)	28.29	33.88	15.93	54.31	68.22	31.29	68.37	66.01	56.78	74.07	65.42	54.24
Expenses (bp)	44.61	19.33	48.00	50.68	18.64	49.00	60.99	12.84	49.00	64.18	15.00	79.00
Creation unit (# of shares)	59,297	20,919	50,000	68,997	29,118	50,000	107,480	74,437	100,000	73,150	24,935	50,000
Creation fees ($)	3,463	3,475	2,200	4,993	4,374	3,500	2,193	2,034	2,300	1,247	250	1,000
Observations		26,536			22,655			7,720			2,781	

Panel B: Country	China			Japan			Brazil			India		
	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median	Mean	Std dev	Median
Age (years)	4.57	2.54	4.10	6.59	5.95	5.75	6.34	4.71	5.07	4.11	2.22	2.76
Market cap ($millions)	337	864	31	1,851	5,142	67	963	2,304	6	377	407	111
Daily volume ($millions)	1.44	6.58	0.01	3.87	12.65	0.00	2.37	5.90	0.02	0.93	1.73	0.24
Daily turnover (%)	0.09	0.19	0.02	0.06	0.13	0.01	0.14	0.24	0.04	0.13	0.15	0.07
Bid-ask spread (bp)	47.15	60.87	21.78	30.50	43.61	15.33	53.94	57.90	37.38	26.20	35.59	12.00
Expenses (bp)	67.89	6.60	65.00	45.27	14.04	48.00	70.64	8.55	69.00	85.91	4.81	86.00
Creation unit (# of shares)	62,372	29,351	50,000	146,078	151,528	100,000	50,000	0	50,000	72,256	53,323	50,000
Creation fees ($)	2,587	1,952	1,900	3,596	1,862	3,500	2,083	715	1,900	2,167	1,342	2,000
Observations		33,260			17,373			10,206			10,003	

Table 3: International Equity ETFs Share Creation and Redemption

The table presents statistics on ETF share creation and redemption. Columns (1) and (2) show the average percentage of trading days when ETF shares were either created or redeemed by authorized participants transacting directly with the ETF. The next columns show the median number of shares in each transaction, and the dollar value corresponding to it, conditional on a creation/redemption transaction taking place. The last columns show the mean size of the transaction relative to the median daily trading volume that month. The mean/median is computed first within a fund and then as another mean/median across funds. The standard deviations are in parentheses. The time period is from January 3, 2012 to December 29, 2017.

Panel A: Development level	% of days				# of shares		Value ($)		% of volume	
	Creation		Redemption		Creation	Redemption	Creation	Redemption	Creation	Redemption
Blended Development	2.29	(14.97)	1.22	(10.98)	200,000	198,500	6.32	4.85	9,062	13,575
Developed Markets	2.13	(14.45)	1.18	(10.82)	300,000	250,000	11.53	7.94	17,648	12,533
Emerging Markets	1.62	(12.61)	1.28	(11.26)	250,000	250,000	8.73	6.82	11,845	11,721
All funds	2.02	(14.05)	1.23	(10.99)	250,000	200,000	8.57	6.34	12,579	12,524
Panel B: Category										
Broad Equity	2.18	(14.59)	1.26	(11.14)	400,000	300,000	13.02	9.66	11,814	9,552
Sector	1.85	(13.49)	1.41	(11.80)	200,000	150,000	5.40	4.83	9,867	14,483
Size and Style	1.64	(12.70)	0.99	(9.91)	250,000	200,000	9.63	5.78	22,890	18,215
Strategy	2.24	(14.80)	1.03	(10.12)	200,000	150,000	5.72	3.77	12,353	17,144
Observations	579,442		579,442		11,436	6,985	11,436	6,985	11,433	6,985

Table 4: ETF Price Premium/Discount Relative to NAV

The table shows the number of ETFs and their market capitalization within each investment category. For the ETFs with available data on net asset values, the table shows the equal-weighted average premium (or discount) of the ETF price relative to its NAV, as well as the equal-weighted and value-weighted time-series volatility of the premium. The market price is taken as the bid-ask average at the end of each trading day. The bid-ask spread is the cross-sectional average of the time-series median bid-ask spread for each ETF. The time period is from January 3, 2012 to December 29, 2017.

Panel A: Focus	Number of funds	Market cap ($m)			Premium (bp)	Volatility of premium		Bid-ask spread (bp)	
		Mean	St dev	Median	Mean	EW	VW	EW	VW
Basic Materials	32	779	3,659	22	-9	55	53	73	67
Broad Equity	250	1279	5300	53	15	41	40	39	36
Consumer Discret	13	113	230	36	-3	18	16	65	61
Consumer Staples	5	123	250	5	12	40	30	103	106
Energy	16	164	350	20	-6	17	18	62	67
Financials	12	93	155	6	8	20	19	77	69
Health Care	10	204	385	7	-1	33	31	89	69
High Dividend Yield	44	271	636	44	13	15	13	38	38
Industrials	10	37	66	14	-8	22	24	70	54
Large Cap	49	354	898	26	15	35	36	40	40
Mid Cap	4	51	67	28	7	32	22	72	56
Natural Resource	20	444	762	94	-2	13	13	37	29
Small Cap	38	349	788	40	7	21	19	59	42
Technology	24	167	279	15	8	21	20	48	50
Telecommunications	4	114	202	18	-3	21	16	90	96
Theme	47	71	144	30	1	43	42	49	47
Utilities	6	45	74	18	-14	24	26	86	96
Total	584	4,658	3,623	476	9	36	35	48	44

Panel B: Quintiles of Trading Volume									
Low		5	11	3	10	49	52	89	89
2		16	19	10	9	48	45	64	59
3		47	55	35	9	30	27	50	46
4		260	306	160	9	22	20	28	22
High		3,438	7,889	1,006	6	22	20	12	7
Panel C: Quintiles of Amihud Illiquidity									
Low		3,309	7,787	975	5	22	20	9	7
2		290	694	136	8	23	20	28	22
3		51	61	36	10	28	28	45	43
4		14	23	8	9	52	49	67	58
High		9	12	4	15	45	49	95	95

Table 5: ETF Premium/Discount by Region and Country

The table presents the number of ETFs and their market capitalization by region and country. The table reports the mean, standard deviation and median of the ETF premium (or discount) relative to its NAV. The market price is calculated as the bid-ask average at the end of each trading day. The bid-ask spread is the cross-sectional average of the time-series median bid-ask spread for each ETF. Others represents countries with only one ETF. The time period is from January 2012 to December 2017.

Panel A: Regional ETFs		Market cap ($m)			Premium (bp)			Bid-ask spread (bp)		
	Number	Mean	St dev	Median	Mean	St dev	Median	Mean	St dev	Median
Asia-Pacific	23	355	777	32	18.41	35.27	10.71	75.88	71.94	47.52
Developed Markets	33	1,218	5,501	40	15.39	30.67	16.69	49.93	37.93	40.79
EAFE	12	6,050	16,151	141	20.86	30.53	17.03	31.17	38.50	10.59
Emerging Markets	80	1,350	7,519	12	17.59	44.58	9.60	58.90	50.69	48.92
Europe	33	775	1,984	31	16.75	18.49	14.11	34.12	25.89	25.65
Global Markets	242	323	1,215	29	7.63	35.52	4.92	51.52	54.94	33.94
Latin America	7	263	600	13	2.21	28.35	-2.86	83.54	55.76	70.38
Middle-East Africa	2	72	4	72	-30.03	23.76	-30.03	74.88	23.46	74.88
Total	432	774	4,626	29	11.41	35.97	8.97	52.79	52.50	36.47

Panel B: Country ETFs	Number	Mean	St dev	Median	Mean	St dev	Median	Mean	St dev	Median
Australia	6	300	704	10	8.03	15.19	8.50	43.60	48.20	26.56
Brazil	8	875	2,364	21	-14.77	17.96	-12.05	59.02	44.23	57.95
China	31	270	783	10.	0.23	33.68	-3.94	50.88	44.94	29.52
Colombia	3	37	50	14	34.62	32.10	31.15	75.46	51.23	98.12
Germany	5	8,245	1,730	13	3.68	6.96	5.05	36.93	44.21	25.90
Hong Kong	3	638	1101	4	11.53	10.09	11.91	58.34	47.89	72.62
India	7	368	436	111	9.41	16.16	11.24	27.68	24.58	20.05
Indonesia	3	288	252	371	-15.13	10.71	-16.05	67.50	99.90	12.29
Israel	2	71	66	71	2.11	9.16	2.11	16.85	8.07	16.85
Italy	2	547	771	547	16.84	18.44	16.83	16.50	12.26	16.50
Japan	20	1,155	4,244	27	-3.96	61.29	6.70	36.13	38.39	26.11
Korea	5	639	1,412	5	1.65	16.69	-1.83	58.66	55.20	39.54
Mexico	2	563	793	563	7.70	15.94	7.70	17.14	20.67	17.13
Poland	2	91	102	91	0.08	5.35	.08	20.21	15.96	20.21
Russia	6	274	420	135	90.02	149.90	38.31	62.84	101.62	20.89
Singapore	2	315	427	315	-2.20	0.30	-2.19	36.98	41.79	36.98
Spain	3	99	160	12	9.10	8.22	4.64	13.45	10.35	11.10
Switzerland	2	272	379	272	7.97	2.27	7.97	16.94	17.95	16.94
Taiwan	4	670	1,335	3	6.34	11.91	5.11	56.12	38.44	60.19
United Kingdom	4	2,035	2,230	1,893	8.29	43.11	27.20	18.60	17.40	14.28
Others	32	265	265	177	4.04	19.58	4.52	28.78	35.30	11.40
Total	152	534	1,849	36	5.82	46.01	4.07	41.79	45.14	24.97

Table 6: ETF Premium and Liquidity

The dependent variable is the absolute value of the premium/discount (in basis points) of the ETF price (closing bid-ask midpoint) over the NAV. The definitions of the explanatory and control variables are in Table A1 in the Appendix. All specifications include fund- year fixed effects. The p-statistics (in parenthesis) are based on double-clustered standard errors across funds and time. The time period is from January 2012 to December 2017. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
FIL	0.0185***	0.0183***	0.0198***	0.0199***
	(0.005)	(0.006)	(0.003)	(0.003)
AIL	-0.00246	-0.00310		
	(0.434)	(0.294)		
FIL/AIL			-0.00684**	-0.00812**
			(0.012)	(0.011)
Size	-0.0492***	-0.0527***	-0.0491***	-0.0526***
	(0.001)	(0.000)	(0.001)	(0.000)
Age	-0.0146	-0.00906	-0.0147	0.00909
	(0.216)	(0.211)	(0.256)	(0.371)
Expense	-0.00489*	-0.00531**	-0.00488*	-0.00532**
	(0.059)	(0.035)	(0.060)	(0.035)
NYSE	-0.00310***	-0.00357***	-0.00313***	-0.00361***
	(0.000)	(0.000)	(0.000)	(0.000)
VIX	0.00196***	0.00172***	0.00196***	0.00173***
	(0.000)	(0.000)	(0.000)	(0.000)
TED	-0.0584	0.0811	-0.0674	0.0771
	(0.372)	(0.237)	(0.302)	(0.258)
Aggregate liquidity	-0.00795***	-0.00431***	-0.00801***	-0.00435***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	23,309	23,309	23,309	23,309
Adjusted R-squared	0.076	0.063	0.076	0.063

Table 7: ETF Premium and Liquidity

The dependent variable is the absolute value of the premium/discount (in basis points) of the ETF price (closing bid-ask midpoint) over the NAV. The definitions of the explanatory and control variables are in Table A1 in the Appendix. All specifications include fund-year fixed effects. The p-statistics (in parenthesis) are based on double-clustered standard errors across funds and time. The time period is from January 2012 to December 2017. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Developed Markets		Emerging Markets		Developed and Emerging	
	(1)	(2)	(3)	(4)	(5)	(6)
FIL	0.00501	0.00322	0.0210**	0.0225*	0.0183*	0.0193**
	(0.646)	(0.767)	(0.042)	(0.053)	(0.098)	(0.041)
AIL	-0.00506		-0.00576		-0.000736*	
	(0.440)		(0.181)		(0.092)	
FIL/AIL		-0.00422*		-0.00844***		-0.00128**
		(0.062)		(0.006)		(0.012)
FIL*Emerging					0.00213*	0.00150*
					(0.059)	(0.056)
AIL*Emerging					-0.00193	
					(0.199)	
(FIL/AIL)*Emerging						-0.00756***
						(0.006)
Size	-0.0587***	-0.0585***	-0.0254	-0.0249	-0.0492***	-0.0497***
	(0.005)	(0.005)	(0.324)	(0.333)	(0.002)	(0.002)
Age	0.00102**	0.000993**	0.000840	0.000873	0.00105***	0.00107***
	(0.018)	(0.021)	(0.120)	(0.104)	(0.002)	(0.001)
Expense	0.00433	0.00410	-0.00943***	-0.00935***	-0.00501	-0.00504
	(0.529)	(0.551)	(0.005)	(0.005)	(0.105)	(0.100)
NYSE	-0.00367***	-0.00367***	-0.00535***	-0.00534***	-0.00438***	-0.00436***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
VIX	0.00110***	0.00110***	0.00295***	0.00296***	0.00198***	0.00199***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TED	-0.00186**	-0.00182**	0.00374***	0.00368***	0.000658	0.000653
	(0.021)	(0.024)	(0.005)	(0.006)	(0.379)	(0.381)
Aggregate liquidity	0.0000849	0.0000990	-0.00481*	-0.00483*	-0.00215	-0.00214
	(0.954)	(0.946)	(0.051)	(0.050)	(0.122)	(0.124)
Observations	8,019	8,019	7,240	7,240	15,259	15,259
Adjusted R-squared	0.066	0.066	0.074	0.074	0.066	0.066

Table 8: Fund Flow and Liquidity

The dependent variable is the monthly net fund flow (in percent). The definitions of the explanatory and control variables are in Table A1 in the Appendix. All specifications include fund-year fixed effects. The p-statistics (in parenthesis) are based on double-clustered standard errors across funds and time. The time period is from January 2012 to December 2017. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Levels for illiquidity measures		AR(1) residuals for illiquidity measures	
	(1)	(2)	(3)	(4)
Lag FIL	-0.000744***	-0.000806***	-0.000828***	-0.000733**
	(0.005)	(0.006)	(0.003)	(0.019)
Lag AIL	0.00000373		-0.00000175	
	(0.531)		(0.779)	
Lag FIL/AIL		0.0000627		-0.0000980
		(0.606)		(0.461)
Size	-0.217***	-0.217***	-0.213***	-0.214***
	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.210**	-0.212**	-0.224***	-0.224***
	(0.010)	(0.010)	(0.007)	(0.007)
Expense	-0.0597***	-0.0598***	-0.0602***	-0.0601***
	(0.000)	(0.000)	(0.000)	(0.000)
NYSE	0.0999	0.0985	0.102	0.102
	(0.735)	(0.739)	(0.730)	(0.731)
Return (past 3-months)	0.795***	0.795***	0.799***	0.799***
	(0.000)	(0.000)	(0.000)	(0.000)
VIX	0.623**	0.622**	0.624**	0.626**
	(0.017)	(0.017)	(0.017)	(0.017)
TED	0.0147	0.0149	0.0141	0.0140
	(0.845)	(0.843)	(0.852)	(0.852)
Aggregate liquidity	-0.0986	-0.0991	-0.0983	-0.0986
	(0.571)	(0.569)	(0.576)	(0.575)
Observations	20,633	20,624	20,416	20,407
Adjusted R-squared	0.015	0.015	0.015	0.015

Table 9: ETF Price Convergence and Liquidity

The table presents the results from the Cox proportional hazard model specified by equations (4) and (5). The dependent variable is the hazard ratio, the conditional probability of the price convergence. The definitions of the explanatory and control variables are in Table A1 in the appendix. All specifications include fund as well as year fixed effects. The p-statistics (in parenthesis) are based on double-clustered standard errors across funds and time. The time period is from January 2012 to December 2017. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
FIL	-0.000829***	-0.000784***	-0.000948***	-0.000903***
	(0.000)	(0.000)	(0.000)	(0.000)
AIL	-0.00000918**	-0.00000779*		
	(0.030)	(0.066)		
FIL/AIL			0.000104	0.000101
			(0.110)	(0.131)
Size	0.0827***	0.0831***	0.0835***	0.0838***
	(0.000)	(0.000)	(0.000)	(0.000)
Age	0.179***	0.163***	0.178***	0.162***
	(0.000)	(0.000)	(0.000)	(0.000)
Expense	-0.00463***	-0.00456***	-0.00459***	-0.00451***
	(0.000)	(0.000)	(0.000)	(0.000)
NYSE	-0.237***	-0.199***	-0.239***	-0.200***
	(0.000)	(0.000)	(0.000)	(0.000)
VIX		-0.0240***		-0.0242***
		(0.000)		(0.000)
TED		0.00299***		0.00292***
		(0.001)		(0.001)
Aggregate liquidity		-0.00500***		-0.00499***
		(0.004)		(0.004)
Observations	24,099	22,210		22,198

Table 10: Liquidity Effect on the Sign Change of the ETF Premium/Discount

The table presents the results from the Cox proportional hazard model. The dependent variable is the hazard ratio, the conditional probability that the ETF premium/discount changes sign. The definitions of the explanatory and control variables are in Table A1 in the appendix. Liquidity mismatch is the difference between the bid-ask spread of the ETF and its underlying portfolio. All specifications include fund-year fixed effects. The p-statistics (in parenthesis) are based on double-clustered standard errors across funds and time. The time period is from January 2012 to December 2017. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
ETF bid-ask spread	-0.000422**	0.000397**	-0.000182*	-0.000261***
	(0.017)	(0.015)	(0.061)	(0.007)
Portfolio bid-ask spread	0.00011	0.000361		
	(0.378)	(0.443)		
Liquidity mismatch			0.00037***	0.000418***
			(0.000)	(0.000)
Size	0.00745*	0.00112**	0.00757	0.00835**
	(0.100)	(0.015)	(0.816)	(0.012)
Age	0.000327***	0.000417***	0.000305***	0.000315***
	(0.002)	(0.000)	(0.000)	(0.000)
Expenses	-0.0083***	-0.0087***	-0.00678***	-0.00658***
	(0.000)	(0.000)	(0.000)	(0.000)
NYSE	0.000385***	0.000196**	0.000228***	0.00124
	(0.000)	(0.023)	(0.003)	(0.116)
VIX		-0.00808***		-0.000517***
		(0.000)		(0.000)
TED		0.0712***		0.0692***
		(0.000)		(0.000)
Aggregate liquidity		0.00296**		0.00298**
		(0.039)		(0.010)
Observations	22,044	20,235	22,013	20,223

Table A1: ETF Distribution

The table reports the distribution, assets under management and fund flows for all ETFs in the ETF Global database by year and asset class for U.S. vs non-U.S. underlying assets. The time period is January 3, 2012 to December 29, 2017.

Panel A: Number of Exchange Traded Funds										
Year	Commodities	Currencies	Equity		Fixed Income		Multi-asset		Real Estate	
			Non-U.S.	U.S.	Non-U.S.	U.S.	Non-U.S.	U.S.	Non-U.S.	U.S.
2012	130	33	342	427	35	145	27	22	11	20
2013	152	39	461	514	60	189	55	27	13	24
2014	158	39	550	554	75	228	66	27	16	24
2015	153	39	641	651	76	238	63	31	22	27
2016	153	35	702	707	86	247	64	41	21	27
2017	158	34	727	790	80	276	78	43	18	28
Total	904	219	3,423	3,643	412	1,323	353	191	101	150

Panel B: Total Assets under Management (Million USD)										
2012	$109,010	$4,341	$245,752	$472,934	$31,498	$142,758	$6,409	$1,760	$3,988	$17,891
2013	$91,857	$4,513	$370,563	$764,910	$24,768	$205,312	$7,948	$2,907	$5,731	$30,158
2014	$63,985	$3,552	$404,031	$975,781	$47,299	$236,847	$6,162	$3,316	$8,557	$35,853
2015	$55,094	$3,964	$462,994	$1,223,631	$29,956	$283,125	$7,345	$3,763	$11,110	$40,817
2016	$68,197	$3,191	$468,637	$1,251,143	$37,093	$352,308	$7,650	$5,580	$11,577	$48,925
2017	$68,044	$2,828	$651,492	$1,675,852	$51,123	$447,737	$9,512	$5,554	$12,246	$54,426
Total	$456,187	$22,390	$3,386,276	$8,253,506	$280,035	$2,157,857	$55,264	$29,300	$67,166	$283,141

Panel C: Non Zero Net Flows Days										
2012	13.69%	6.47%	9.85%	19.94%	9.44%	17.88%	6.01%	15.78%	6.81%	20.94%
2013	13.62%	6.19%	9.95%	21.27%	9.25%	17.72%	6.24%	17.80%	6.59%	21.12%
2014	13.74%	6.43%	10.23%	21.22%	9.01%	16.82%	6.49%	21.20%	6.51%	21.49%
2015	14.40%	6.75%	9.92%	20.64%	9.58%	16.39%	7.50%	19.23%	6.14%	23.03%
2016	14.90%	7.10%	9.28%	19.07%	9.68%	16.76%	7.77%	18.10%	5.56%	22.61%
2017	14.47%	7.34%	9.45%	18.12%	10.71%	16.12%	7.50%	18.12%	6.41%	20.81%

Table A2: Variable Definitions

Variable	Definition
Panel A: Market characteristics	
ETF premium	ETF market price over its NAV minus one. A number greater (less) than zero represents ETF premium (discount).
Net Fund Flow	The change in shares outstanding times NAV as a percentage of last month's assets under management.
NYSE	A dummy variable equal to one if ETF is traded on NYSE Arca.
BATS	A dummy variable equal to one if ETF is traded on Bats BZX Exchange.
NASDAQ	A dummy variable equal to one if ETF is traded on NASDAQ.
Idiosyncratic volatility	The standard deviation of the residuals of ETF returns regressed on market index returns.
Short Interest	Short interest as a fraction of shares outstanding.
Panel B: Illiquidity measures	
FIL	ETF illiquidity: the absolute daily return to dollar volume aggregated over a month.
AIL	Underlying portfolio illiquidity: the weighted average of the illiquidity measures for all holdings of the ETF portfolio.
FIL/AIL	liquidity mismatch between the ETF and underlying portfolio of securities.
Panel C: Fund characteristics	
Size	The natural logarithm of assets under management (AUM).
Age	The natural logarithm of $(1 + age)$. The age of the fund is in years since inception.
Expense	The expense ratio of the fund in basis points of AUM.
Panel D: Macro variables	
VIX	Changes in the VIX index (volatility implied by S&P 500 index options).
TED spread	The difference between three-month Treasury bill and three-month LIBOR based on US dollars.
Aggregate liquidity	Pastor and Stambaugh (2003) aggregate level of liquidity

A3: Fund Flow and Liquidity: Panel Vector Autoregression

Panel A presents GMM estimates from a panel VAR model with endogeous variables monthly net fund flow (in percent) and ETF Amihud illiquidity measure. The definitions of the explanatory and control variables are in Table A1 in the Appendix. All specifications include fund-year fixed effects. The p-statistics (in parenthesis) are based on double-clustered standard errors across funds and time. Panel B presents the results from Granger causality tests. The time period is from January 2012 to December 2017. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: GMM estimation of panel VAR

	Level of Amihud measure		AR(1) residuals of Amihud measure	
	(1)	(2)	(3)	(4)
	Net Fund Flow (%)	ETF Illiqudity	Net Fund Flow (%)	ETF Illiqudity
Lag Net Fund Flow	-0.180	0.00124	-0.057	0.000993
	(0.364)	(0.783)	(0.211)	(0.572)
Lag ETF illiquidity	-0.097***	0.0166*	-0.330***	0.057**
	(0.004)	(0.063)	(0.000)	(0.011)
Size	-0.0118***	0.493***	-0.0120**	0.551**
	(0.007)	(0.006)	(0.016)	(0.013)
Age	0.00206	-0.722***	0.00245	-0.770***
	(0.739)	(0.003)	(0.725)	(0.009)
Expense	-0.193	16.44	-0.201	18.32
	(0.313)	(0.112)	(0.354)	(0.136)
NYSE	0.00146	-1.060	0.00219	-1.195
	(0.913)	(0.128)	(0.882)	(0.149)
Return (past 3-months)	0.0207***	-0.594***	0.0209***	-0.619***
	(0.000)	(0.000)	(0.000)	(0.001)
VIX	0.000866***	-0.0373***	0.000854***	-0.0358***
	(0.003)	(0.000)	(0.004)	(0.000)
TED	-0.000224	0.221**	-0.000188	0.254**
	(0.930)	(0.018)	(0.948)	(0.026)
Aggregate liquidity	-0.00114	-0.0106	-0.00106	-0.0251
	(0.245)	(0.690)	(0.286)	(0.422)
Observations	20,572		20,406	

Panel B: Panel VAR-Granger causality Wald test

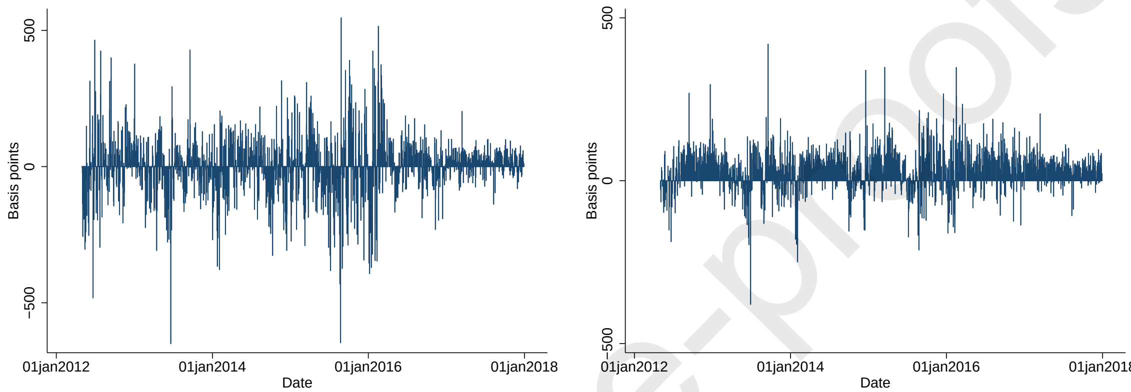
Ho: Excluded variable does not Granger-cause Equation variable

Ha: Excluded variable Granger-causes Equation variable

Equation	Excluded	χ^2	p-value
Net fund flow	ETF illiquidity	8.516	0.004
ETF illiquidity	Net fund flow	0.076	0.783

Figure 1: Premium/Discount for ETFs with Similar Underlying Portfolios

Panel A: Emerging Market ETFs: EEM (left) and SCHE (right)



Panel B: Developed Market ETFs: EFA (left) and SCHF (right)

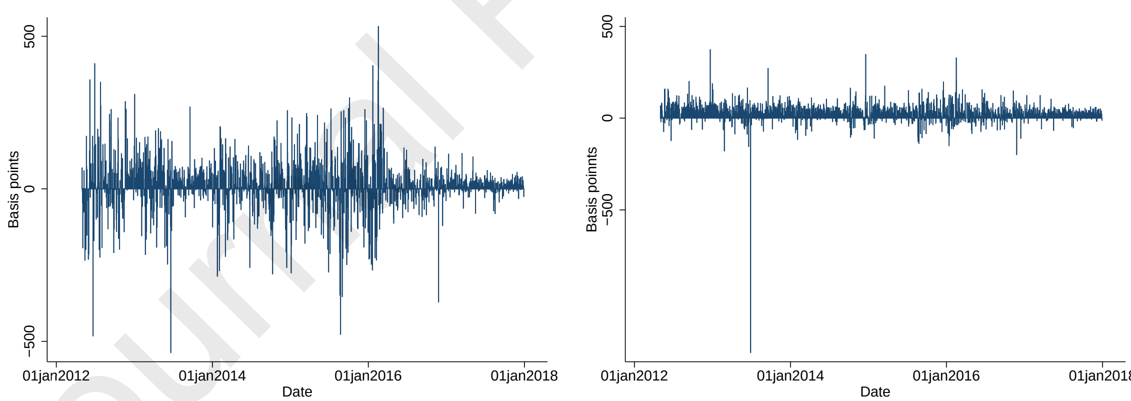
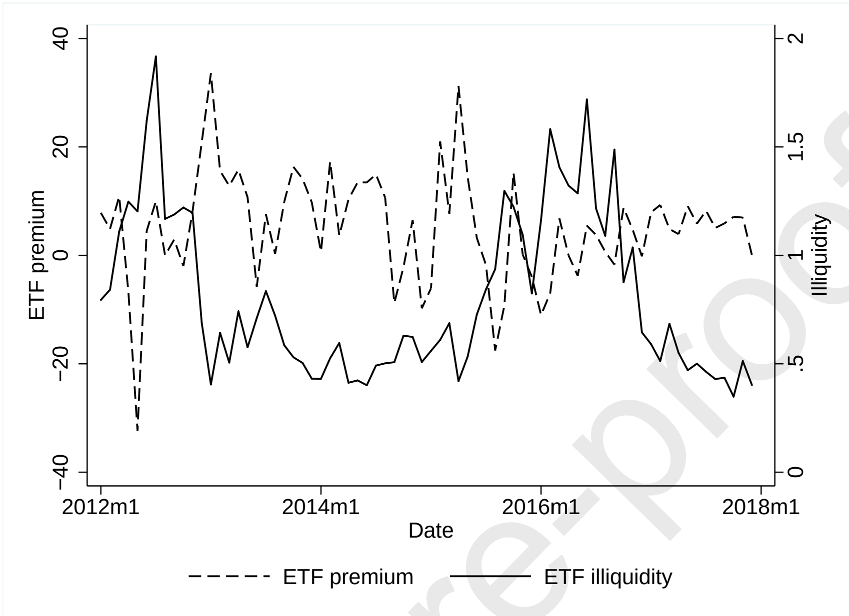


Figure 2: ETF Premium/Discount and Illiquidity

Panel A: ETF Premium/Discount and ETF Illiquidity



Panel B: ETF Premium/Discount and Portfolio Illiquidity