Financial Analysts Journal | A Publication of CFA Institute

Inefficiencies in the Pricing of Exchange-Traded Funds

Antti Petajisto

Antti Petajisto is portfolio manager at LMR Partners, London.

The prices of exchange-traded funds (ETFs) can deviate significantly from their net asset values (NAVs), in spite of the arbitrage mechanism that allows authorized participants to create and redeem shares for the underlying portfolios. The deviations, typically within a band of about 200 bps, are larger in funds holding international or illiquid securities. To control for stale pricing of the underlying assets, I introduce a novel approach that uses the cross section of prices on a group of similar ETFs. The average pricing band remains economically significant at about 100 bps, with even larger mispricings in some asset classes. Active trading strategies exploiting such inefficiencies produce substantial abnormal returns before transaction costs, providing further proof of short-term mean reversion in ETF prices.

Disclosure: The first draft of this article was written when the author was a finance professor at the NYU Stern School of Business.

CE Credits: 1

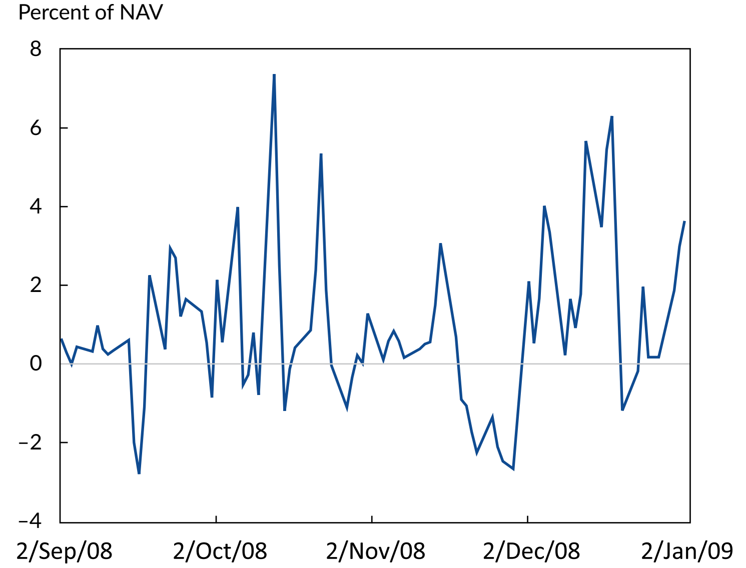
In late 2008, some investors spotted an attractive opportunity to buy municipal (muni) bonds or junk bonds, because both types of bonds had declined dramatically in value in the immediate aftermath of the Lehman Brothers collapse, arguably more than could be justified by realistic expectations of future defaults. Exchange-traded funds (ETFs) provided a convenient new vehicle for investors to take advantage of this opportunity: Any investor, including a retail investor, could easily buy a diversified portfolio of muni bonds or junk bonds, pay a low expense ratio, trade positions intraday, and use leverage to capitalize more aggressively on the opportunity. Efficiently executing such trades with ETFs, however, turned out to be more complicated.

Two muni bond funds (tickers: PZA and MLN) both offered exposures to very similar portfolios of long-term muni bonds diversified across the entire United States. But with respect to their pricing, one was sometimes trading at a premium relative to its net asset value (NAV) whereas the other was trading at a discount at exactly the same time. The difference in the end-of-day premiums between the two funds varied from about +7% to –3% of NAV (**Figure 1**). Junk bond ETFs exhibited similar behavior, with the difference in premiums between the two largest funds (tickers: JNK and HYG) varying from +7% to –11% of NAV. The differences in premiums were driven by mean-reverting shocks to prices; that is, it was enormously important for an investor to correctly pick which fund to trade each day. Such pricing behavior was particularly surprising in the junk bond ETFs, given that each fund had several billion dollars

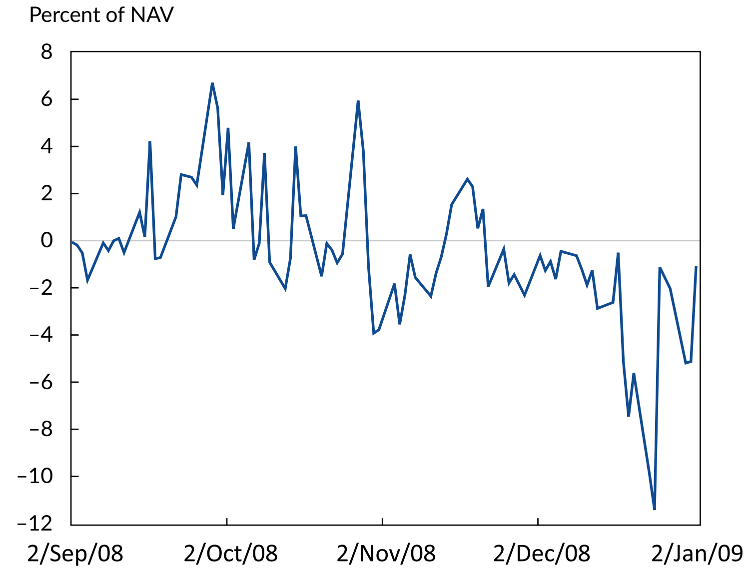
I thank Ned Elton, Rob Engle, Tom Faust, Ross Garon, Gary Gastineau, Marcin Kacperczyk, Juhani Linnainmaa, and Albert Menkveld for helpful discussions; David Smalling for early research assistance; and seminar participants at NYU Stern, Temple University, BlackRock, and the Society of Quantitative Analysts. I also thank SAC Capital and Morningstar for providing data for this study. I gratefully acknowledge the financial support of INQUIRE Europe.

Figure 1. Premiums on Similar ETFs

A. Premium on PZA vs. MLN



B. Premium on JNK vs. HYG



Notes: The top panel shows the end-of-day price premiums on two similar muni bond funds (PZA and MLN) relative to their NAVs as well as the difference in the two premiums over time. The bottom panel shows the difference in premiums for two similar junk bond ETFs (JNK and HYG).

in assets and traded large volumes every day with tight bid–ask spreads.

Although the magnitudes in these examples are unusually large,[1] this article provides empirical evidence that smaller inefficiencies in the pricing of ETFs are not limited to a few funds or a particular period. In fact, the cross section of ETFs routinely exhibits some economically significant differences between the ETF share price and the value of the underlying portfolio, especially in some asset classes, indicating that the unsophisticated investor may face an unexpected additional cost when trading ETFs.

Given the lack of attention that ETF premiums have received from investors, many of them appear implicitly to assume that ETF prices stay extremely close to NAVs. This assumption may be understandable because of the arbitrage mechanism that exists for ETFs: If the price is below the NAV, an arbitrageur can purchase ETF shares, redeem them for the underlying assets held in the ETF portfolio, and then sell the underlying assets at their prevailing market prices, which add up to the fund's NAV.[2] If the ETF price is higher than the NAV, an arbitrageur can do the reverse and create new ETF shares from the underlying assets. This action generates a pure arbitrage profit, minus the transaction costs of buying or selling the underlying portfolio. The efficiency of ETF prices would thus be expected to depend on transaction costs and any other limits to arbitrage that might deter arbitrageurs from trying to profit from a mispricing.

Another issue complicating the arbitrage trade is that the officially published NAV may not fully reflect the current value of the ETF portfolio because of stale pricing. NAV is computed on the basis of the latest closing prices of the underlying securities or the latest bid prices in fixed-income markets (e.g., Gastineau 2010; Tucker and Laipply 2010). This computation can be a problem for illiquid securities, such as high-yield bonds, or for securities traded in international markets, such as Japan, where the trading day ends before it even begins in the United States. Hence, estimating the true NAV, as distinct from the published NAV, becomes a more complicated task. Furthermore, in the case of international securities traded in different time zones, it may even be impossible to enter into simultaneous offsetting transactions involving ETF shares and the underlying portfolio. These concerns are likely to reduce the effectiveness of arbitrage and to allow greater mispricings in ETFs.

In my study, a key research goal was to provide comprehensive evidence of the magnitude of ETF premiums across all US-listed ETFs and all underlying asset classes and over time. An essential part of this objective was to determine the magnitude of the "true" ETF premiums when stale NAVs are accounted for. Moreover, I wanted to understand what the premiums depend on and how investors should trade ETFs to avoid being adversely affected by the premiums.

I started by computing the premiums (positive or negative) of ETF midquote prices relative to NAVs for all categories of funds traded in the US market. I also documented the time-series evolution of the premiums to see whether the efficiency of the market has changed over time. Most importantly, I devised a novel approach to addressing the stale-pricing issue: I sorted funds into groups with nearly identical underlying portfolios, using the average market price of the group as a real-time proxy for the true underlying value of the funds. Any cross-sectional dispersion of an ETF price around its group mean is thus likely to be explained by mispricing rather than by stale NAVs. Owing to the dramatic growth of the ETF sector, I focused mostly on the data over January 2007–December 2014, because older data may be a poor guide to the present situation in the ETF marketplace.

I found that the average premium across all funds is only 6 bps; so, on average, ETFs are neither underpriced nor overpriced, in contrast to closed-end funds, in which the absence of share creation and redemption allows some funds to trade 10%–20% or more below their NAVs. But the volatility of the ETF premium is nontrivial at 49 bps, which means that with 95% probability, a fund is trading at a premium between about –96 bps and +96 bps, or within a 192 bp band. The value-weighted numbers are only slightly smaller. This range is certainly economically significant and a potential source of concern for an ETF investor. Diversified US equities and US government bonds exhibit less volatile premiums, whereas international equities and municipal and corporate bonds exhibit volatilities of 40–140 bps around NAVs.

Although stale pricing explains part of the premiums, it is certainly not a complete explanation. When I computed the volatility of the premium relative to the mean of a group of similar ETFs, I found that this peer-group-adjusted volatility is still 26 bps. Thus, even after adjusting for stale pricing, the average ETF trades 95% of the time at a premium between –52 bps and +52 bps, or within a band of 104 bps. The numbers are, again, higher for some fund categories and twice as high for some international funds.

Bid–ask spreads cannot explain these premiums. I used the bid–ask midpoint price to measure all premiums, but spreads are relatively tight for commonly traded ETFs and thus these results are robust to using any price within the bid–ask spread.

The dollar amounts are also significant. Because of the large trading volume of ETFs, the historical premiums in actual ETF transactions amount to over $40 billion a year relative to NAVs and roughly $20 billion when adjusted for stale NAVs—a rather large amount of money to pay for trading at inefficient prices, suggesting that investors should not ignore these effects when trading ETFs.

Furthermore, an active trading strategy built to exploit these cross-sectional differences in ETF premiums generates attractive profits: For a simple unlevered long–short strategy trading at midquote prices (thus neither paying nor earning the bid–ask spread), the historical Carhart alpha is around 7% a year and the annualized information ratio is 4.8 using all the ETFs, with the alpha rising to 16% a year using only the categories most prone to mispricing. This example provides another illustration of the inefficiencies that remain in the ETF market and the potential pitfalls for the average ETF investor.

Positive premiums on ETFs lead to more share creation—and vice versa for negative premiums—indicating that arbitrageurs are actively using the ETF share creation and redemption process to trade against these mispricings. Once new shares are created, there is downward price pressure on the same day and the subsequent two days, which in turn pushes positive premiums back toward zero.

The cross-sectional dispersion in ETF premiums peaked during the financial crisis in late 2008, and it has remained at a nontrivial level since then. It is correlated with the VIX index and the Treasury–Eurodollar (TED) spread, which are proxies for the availability of arbitrage capital, as well as the average closed-end fund discount, which can proxy for investor sentiment.

There is relatively little prior literature on the efficiency of ETFs as investment vehicles. Early studies include Ackert and Tian (2000, 2008); Elton, Gruber, Comer, and Li (2002); Poterba and Shoven (2002); and Delcoure and Zhong (2007). Marshall, Nguyen, and Visaltanachoti (2010) looked at the efficiency of intraday pricing of three S&P 500 Index funds. In contrast, Wurgler (2011), Sullivan and Xiong (2012), and Ben-David, Franzoni, and Moussawi (2014) discussed the impact of ETF-induced trading on the efficiency of the underlying asset markets. The study perhaps closest to mine is Engle and Sarkar (2006), who analyzed similar questions about ETF premiums and tried to adjust for stale NAVs with an econometric model. However, all these studies examined only a handful of ETFs. For example, Engle and Sarkar (2006) used a sample of 37 ETFs ending in 2000, and Delcoure and Zhong (2007) used 20 ETFs ending in 2002. Since then, the size of the ETF market has increased to over 1,400 live US-traded ETFs, with about $2 trillion in assets at the end of 2014.[3] One exception is Ben-David et al. (2014), who used a broad cross section of ETFs but focused on a different research question and did not adjust for stale NAVs, which is a key issue when investigating ETF premiums.[4]

In this article, I offer the first comprehensive study of the pricing efficiency of all US-listed ETFs after the dramatic surge in new products. To accomplish this objective, I present a number of descriptive summary statistics for today's ETF market to provide the necessary context for researchers and investors. Very importantly, I also propose a new and generally applicable methodology for dealing with stale NAVs without having to make any assumptions about the price dynamics of the underlying portfolio. In addition to ETFs, the same methodology could be helpful in the fair-value pricing of closed-end funds or securities trading across different time zones.

Data

In my study, I combined six sources of ETF data. The first source is CRSP, which I used for daily prices and returns. CRSP covers all US-listed live and dead ETFs, including commodity funds but not exchange-traded notes (ETNs). The second source is Bloomberg, which covers daily NAV data for essentially all live funds as of April 2010 (the time of my first Bloomberg snapshot) or later, going back to the inception of each fund. From 1995, the Bloomberg data include about 60%–100% of all ETFs and 90%–100% of all ETF assets. The third source is iShares, covering daily NAV data for iShares funds from inception to July 2009. The iShares funds account for over 50% of all ETFs until the end of 2005 and generally account for about half of all assets since the beginning of 2005. The NAV data for any remaining funds are from OpenTick, a defunct data vendor that used to provide minute-by-minute estimates of NAVs throughout the trading day for a cross section of funds from all fund families. The OpenTick data cover about 40%–50% of funds and 30%–50% of fund assets between October 2006 and February 2009. Collectively, the three sources of NAV data cover about 87%–99% of ETFs and 97%–100% of ETF assets over 2007–2014, the focus of most of my analysis. I selected data from Bloomberg, iShares, and OpenTick in that order, and thus the overwhelming majority of NAV data points are from Bloomberg (to the extent that Bloomberg alone would still be a reasonable source of NAV data).

The fifth data source is Morningstar, which I used for fund names, style categories, and benchmark indexes. Multiple snapshots were downloaded between March 2010 and June 2014 for live ETFs and then merged into one dataset, accounting for 88%–100% of my cumulative fund sample and at least 99.8% of ETF assets over 2007–2014. Survivorship bias was not an issue because I did not study the performance of individual ETFs.

The sixth data source is the consolidated NYSE TAQ database, which was aggregated from individual transactions and quotes into five-minute intervals. I used it for intraday calculations, including bid–ask spreads, prices, and trading volume. I did not use intraday data for the bulk of my analysis; instead, I used such data mostly as a robustness check to confirm that daily observations did not miss some relevant intraday behavior.

To mitigate concerns about illiquidity in the shares of smaller ETFs, I computed the end-of-day price as the average of the bid price and the ask price at market close instead of using the official closing price (i.e., the latest transaction price). I also computed all ETF returns from the bid–ask midpoint (plus dividends) rather than using the returns from CRSP, following Engle and Sarkar (2006). In some parts of my analysis, I eliminated funds with less than $10 million in assets or less than $100,000 in daily trading volume.

Background on ETFs

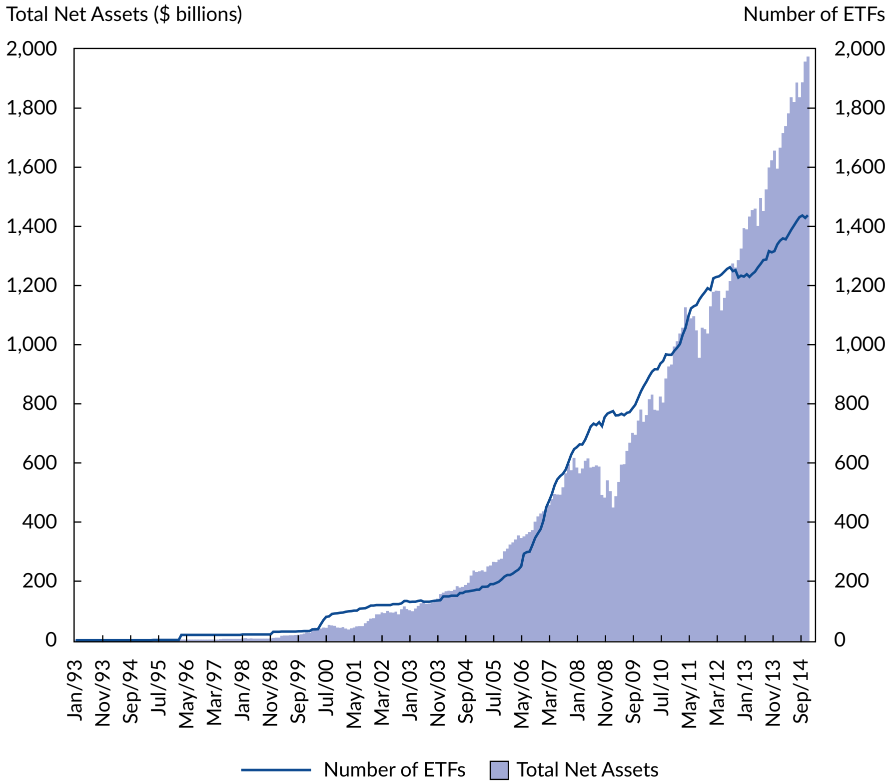
In this section, I provide the necessary context for researchers and investors by examining the dramatic growth of the ETF sector, cross-sectional ETF characteristics, and share creation and redemption.

Growth of the ETF Sector. Before ETFs, most individual investors were effectively limited to investing in open-end and closed-end mutual funds or directly in stocks. Relative to mutual funds, ETFs advertise several benefits for investors. ETF fees are generally comparable to or even lower than those of the lowest-cost index funds. The ETF structure allows funds to minimize portfolio turnover, thus generating lower trading costs than comparable open-end index mutual funds. ETFs can be more tax efficient. They offer intraday trading, can be sold short or bought on margin, and can all be purchased conveniently on the investor's existing brokerage account. Investors have paid attention, and the ETF sector has risen to become a very serious challenger to the mutual fund industry.

Figure 2 shows the explosive growth of the ETF sector in the last few years. The first US-listed ETF, the SPDR trust (ticker: SPY) from State Street, was launched in January 1993. Three years later, in March 1996, the first competing firm (WEBS, later acquired by iShares) entered the field with 17 international single-country ETFs. The market experienced significant growth, reaching 200 funds in October 2005, and then started an even more explosive growth stage, reaching 1,435 live US-listed funds with $1.97 trillion in assets in December 2014. ETFs were among the few investment vehicles receiving broad inflows throughout the financial crisis in 2008.

Cross-Sectional ETF Characteristics. **Table 1** describes my sample of ETFs in 2014, showing the whole distribution of some key

Figure 2. The Size of the ETF Sector in the United States



Notes: For all ETFs traded in the United States, this figure shows the number of ETFs and their total market cap from the inception of the first ETF (January 1993) to December 2014. ETNs are excluded, but all ETFs, including commodity and currency funds, are included.

characteristics. The median fund has $111 million in assets, but the distribution is heavily skewed in terms of asset size, with the largest fund (SPY) accounting for $216 billion. Dollar trading volume is even more skewed, with the median fund trading $1.2 million a day and the most active fund (SPY) trading $21 billion a day. Relative to a fund's market capitalization, daily trading amounts to about 1.1% for the median fund, implying an annual turnover of around 300%, but the most active funds can trade more than their market cap in a single day. The median ETF closing bid–ask spread is 15 bps, varying from as low as 1 bp for the most liquid funds to several percent for the least liquid funds, reflecting the wide disparity in trading volume across funds. Unlike an individual stock, for which market makers must post a large spread to offset the adverse selection problem (i.e., they may lose money trading with someone who has private information), an ETF is a diversified portfolio on which idiosyncratic information has a smaller valuation impact, and thus ETF spreads should generally be lower than the spread for a stock with similar trading volume.

The median fund generates a 26% annual turnover by its own trading. Some turnover is unavoidable, even for passive funds, because of changes in the underlying index. In particular, funds holding front-month futures positions need to trade often because they roll over their positions regularly, whereas a diversified large-cap equity index requires little turnover if the fund uses only in-kind creations and redemptions. The annual expense ratio of a median fund is 50 bps of net assets, varying from 4 bps to 701 bps across funds.[5]

Table 1 also shows the same statistics for two earlier snapshots—December 2010 and December 2006. The distribution of ETF characteristics was relatively stable over 2006–2014. In 2006, however, the median fund had slightly more assets, greater trading volume, and a slightly lower bid–ask spread, reflecting the recent proliferation of

Table 1. Sample Statistics of ETFs

Variable	Mean	Percentile Min.	5	25	50	75	95	Max.	*N*
2014									
Market cap ($ millions)	1,383	0.1	3.1	20	111	529	5,762	215,916	1,423
Daily volume ($ millions)	46	0.0	0.0	0.2	1.2	6	112	21,113	1,435
Daily turnover (%)	4	0.0	0.3	0.6	1.1	2	15	146	1,432
Bid–ask spread (bps)	28	1	2	7	15	32	88	711	1,435
Fund turnover (%)	47	1	4	11	26	52	146	1,335	1,014
Expenses (bps)	56	4	12	32	50	75	95	701	1,194
2010									
Market cap ($ millions)	1,023	1.2	3.7	20	92	428	4,328	90,965	968
Daily volume ($ millions)	71	0.0	0.1	0.3	1.2	7	183	23,792	978
Daily turnover (%)	6	0.1	0.5	0.9	1.7	4	29	147	977
Bid–ask spread (bps)	19	1	3	8	14	23	57	453	978
Fund turnover (%)	51	1	4	12	29	57	169	1,232	720
Expenses (bps)	55	7	14	35	54	71	95	150	860
2006									
Market cap ($ millions)	1,134	5.0	9.3	29	125	671	4,529	63,725	376
Daily volume ($ millions)	72	0.0	0.1	0.4	1.6	9	112	9,160	376
Daily turnover (%)	4	0.2	0.5	0.9	1.7	4	16	80	350
Bid–ask spread (bps)	14	1	5	10	13	16	23	153	376

Notes: This table shows the cross-sectional distribution of various characteristics of US-listed ETFs for three end-of-year snapshots in 2014, 2010, and 2006. Daily volume and turnover represent the trading by investors in ETF shares, computed as the mean throughout the year. The bid–ask spread of an ETF is computed as the median end-of-day closing spread. Market capitalization is the last available month-end value that year. Fund turnover refers to annual turnover of securities within the ETF's portfolio (thus excluding in-kind creations and redemptions).

ETFs and the testing of investor appetite for new products introduced in various niche categories.

Share Creation and Redemption. To create new ETF shares, an investor must be an "authorized participant"—for example, a broker/dealer who has entered into an agreement with the ETF trustee. Typically, ETF shares are created in "creation units" of 50,000 or 100,000 shares, with dollar values ranging from $300,000 to $10 million. Most creations occur as in-kind transactions: The investor submits a portfolio that matches the specifications given by the fund trustee before the end of the trading day, and new ETF shares are created for the investor at the end of the trading day. The trades are settled three days later.

The authorized participant must pay a fixed-dollar fee, usually $500 to $3,000, for each creation transaction regardless of the number of creation units involved. For SPY, its fixed fee of $3,000 would amount to about 3 bps for a single creation unit worth about $10 million, or 1 bp for three creation units worth about $30 million. The process is similar for share redemptions, with identical fees. These transaction costs, combined with the costs of trading the underlying securities, would thus be expected to set the boundaries of how much the ETF price can diverge from its NAV.

Some ETFs also allow investors to create and redeem creation units in cash. These transactions occur at the fund's published end-of-day NAV, much like purchases of open-end mutual funds whereby the fund must use the new cash to purchase more securities for its underlying portfolio. In ETFs, however, the investor who creates new shares must pay additional fees to cover the transaction costs incurred by the fund. These fees vary widely. For example, ProShares ETFs, based on relatively liquid underlying assets, charge additional fees of only 0–3 bps for cash creations, whereas the iShares high-yield bond fund (ticker: HYG) may charge up to 3% for cash creations and 2% for redemptions, reflecting the higher transaction costs of the underlying assets.[6]

In-kind share creation exposes an arbitrageur to two risks: (1) the timing risk arising from the nonsimultaneous purchase and sale of the ETF shares and the underlying portfolio and (2) unpredictable transaction costs, especially for illiquid assets. Although cash creation eliminates these risks, it can be much more expensive and is not always available. Thus, even if arbitrageurs compete aggressively in this activity, their actions are likely to leave some nontrivial mispricings, at least for the types of ETFs in which the limits to arbitrage are most significant.

Table 2 presents statistics on share creations and redemptions over January 2007–December 2014. I computed the fraction of trading days when each ETF experienced share creations or redemptions and then computed the mean and median across all funds. On average, creations or redemptions occurred on 13% of all trading days. However, this measure is skewed by many small funds with little or no activity, because the median fund had such activity on only 6% of trading days.

The other columns in Table 2 are all conditional on creations or redemptions taking place that day. The median number of shares created or redeemed is 100,000—a common size for one or two creation units—whereas the mean is 257,000 shares. The median dollar value of these transactions is $3.4 million, whereas the mean is $11.2 million. As a fraction of a fund's total assets, the median transaction accounts for 5%, whereas the mean accounts for 22%. Relative to the daily ETF trading volume, these fractions are much larger, with the median and mean creation or redemption transaction accounting for 310% and 3,087% of daily volume, respectively.

Economically, these numbers suggest that the size of a creation unit for a typical ETF is indeed large. Even if an arbitrageur participates in every single trade in a fund and always on the same side, in most funds the arbitrageur would still need several days to accumulate a position large enough to offset the creation or redemption of a single creation unit. This situation introduces timing risk, which makes it harder to arbitrage small mispricings by using the ETF share creation/redemption process—thus making it less surprising if prices do not closely track NAVs for many funds. The fund categories most affected by infrequent creations and redemptions are those with the most difficult-to-trade underlying assets—including international equities (differences in trading hours introduce timing risk even when the underlying assets are actively traded in their own markets) as well as corporate and municipal bonds—whereas funds holding US equities and US government bonds experience more creation and redemption activity, on average.

The bottom two panels of Table 2 show the same statistics across funds sorted into quintiles by market cap and trading volume. The larger and more traded ETFs have much more frequent creations and redemptions. In spite of the larger size of creations for larger funds, such creations account for a much smaller fraction of daily trading volume, making arbitrage activity easier and more frequent in these funds. In contrast, creations for small and less liquid funds are often driven by investor flows, whereby a long-term investor buys the creation basket (which may have more liquidity) and works with the ETF sponsor's capital markets desk to immediately convert the basket into new ETF shares.

ETF Premiums relative to NAVs

In this section, I present my results on ETF premiums relative to the funds' official NAVs as well as

Table 2. Share Creation and Redemption Activity, January 2007–December 2014

	Percent of Days		Shares (thousands)		Value ($ millions)		Percent of Shares		Percent of Volume	
Category	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
US equity: Diversified	14	6	273	100	17.6	4.0	22	3	4,727	383
US equity: Sectors	16	9	244	100	8.9	2.8	15	4	1,987	218
US bonds: Government	15	9	217	100	16.2	5.2	17	3	1,346	255
US bonds: General	15	9	201	100	10.7	4.4	16	3	1,477	317
US bonds: Munis	8	5	124	100	4.6	2.5	12	3	3,584	512
International equity	9	5	306	100	10.7	3.5	22	5	2,934	369
International bonds	9	3	197	100	8.9	5.1	20	5	9,772	461
Allocation	6	3	763	50	15.2	1.7	68	13	13,327	890
Commodities	13	9	349	100	14.2	4.5	18	4	1,436	173
Miscellaneous	12	6	156	50	7.5	3.3	27	11	1,522	227
All	13	6	257	100	11.2	3.4	22	5	3,087	310
Fund size quintile										
Large	35	29	606	200	32.0	10.3	4	1	141	56
4	15	13	228	100	8.7	4.7	8	2	464	174
3	8	6	241	100	7.5	2.8	20	5	2,376	351
2	4	3	102	50	3.7	2.3	29	14	3,740	751
Small	1	1	70	50	2.2	1.5	52	50	9,947	1,559
All	13	6	257	100	11.2	3.4	22	5	3,087	310
Trading volume quintile										
Large	35	29	588	200	30.9	9.6	6	1	135	41
4	14	12	202	100	7.7	4.2	9	2	466	162
3	7	6	168	100	5.5	2.6	19	5	1,747	379
2	3	2	161	50	4.9	2.2	36	17	5,339	829
Small	1	1	83	50	2.7	1.9	52	50	11,092	2,148
All	13	6	257	100	11.2	3.4	22	5	3,087	310

Notes: The first two columns show the 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 number of shares (in thousands) in each transaction and the dollar value corresponding to it, conditional on a creation/redemption transaction taking place. The last columns show the size of the transaction relative to both the total ETF shares outstanding and the average daily trading volume that month. The median is computed first within a fund and then as another median across funds; the mean is similarly computed first within a fund and then across all funds.

on the relationship between ETF premiums and share creations.

Sample and Methodology. In my study, I defined the ETF price premium as the percentage deviation of the ETF price (closing midquote) from the NAV. For simplicity, I called it a "premium" even when it was negative—that is, when the ETF was trading at a discount. I weeded out a handful of premiums greater than 20% in absolute value because they were all the result of data errors, but in general, my data sources seemed relatively clean.[7]

I focused on the premiums in the last eight years of the data—that is, January 2007–December 2014, thus excluding the pre-2007 period—for two reasons. First, owing to the dramatic growth of the ETF industry in the more recent years (Figure 2), this approach was the only way to obtain a broad cross section of funds for the entire period. Second, it would be questionable to assume that the pricing of ETFs has not changed in any way while the industry has undergone such an explosive period of growth, and thus the early years of the ETF market may not be informative about the current state of ETF pricing. **Table 3** reports the ETF sample in the first columns. There is a cumulative total of 1,813 funds (including dead funds) over the sample period, with $1.97 trillion in assets at the last available date for each fund. There are NAV data for 1,670 funds, extending across multiple fund families. The largest categories are equity funds: $905 billion in diversified US equity funds, $273 billion in sector funds, and $380 billion in international funds. Bond funds collectively have about $304 billion. Inverse ETFs are in the "bear market" category, with $13 billion. Commodity funds have $51 billion, mostly in precious metals, particularly gold. The categories are from Morningstar and apply only to live funds, so dead funds are in their own category.[8]

I computed the statistics on the ETF premiums as follows. First, I calculated for each fund the average level of the premium and the time-series volatility of the premium. I then averaged across funds within each category to obtain the average premium and the average volatility of the premium for the category. To compute value-weighted averages, I weighted each fund by its average market capitalization over the period.

Estimates of Premiums. The average premium is only 6 bps, which indicates that the typical ETF is neither underpriced nor overpriced. The only exceptions, with slightly positive average premiums of around 20 bps, are bond funds, whose NAVs are based on bid quotes and thus understated, and international equity funds.[9] However, the time-series volatility of the premium is 49 bps, which suggests that ETF prices fluctuate considerably around NAVs even if the average level is correct. The value-weighted average volatility is comparable at 40 bps, so the result is not limited to smaller funds.

Economically, the equal-weighted volatility tells us that the typical fund is trading within a range of –96 bps to +96 bps around its NAV, with a 95% probability. Given that some funds are competing for cost-conscious investors by shaving a few basis points off their fees to bring them below 10 bps a year, there is a risk that some investors are overlooking a potentially much bigger cost due to an adverse premium on the transaction price. Conversely, transacting at an attractive price can offset the cost of investing in a higher-fee ETF.

The smallest premiums generally exist in diversified US equities, US government bonds, and shortest-maturity bonds. At the other end of the spectrum, international equities, international bonds, and illiquid US-traded securities, such as municipal bonds and high-yield bonds, exhibit volatilities of up to 144 bps, which translates to a 95% confidence interval of almost 6%. That result is qualitatively consistent with the limits-to-arbitrage hypothesis because the securities with the highest transaction costs and the least transparent (and stalest) NAVs have the most volatile premiums. But can these costs really explain the entire magnitude of the premiums? One piece of evidence comes from sector funds. Such sectors as real estate, technology, health care, and financials contain mostly US-focused funds, which have liquid and transparent holdings, and yet these categories have premium volatilities of up to 40 bps, which cannot be explained by stale pricing. For more general evidence, we must deal directly with the stale-pricing issue (discussed later in the article).

Table 3. ETF Price Premiums (and Discounts) relative to NAV, January 2007–December 2014

Category	Market Cap ($ millions)	No. of Funds		Avg. Premium (bps)	Volatility of Premium (bps)			Bid–Ask Spread (bps)	
		All	NAV		EW	VW	VW Min.	EW	VW
US equity: Diversified	904,995	301	296	0	18	9	7	18	3
Large blend	454,381	71	68	1	20	9	7	17	2
Large growth	110,506	39	38	–1	20	8	6	19	3
Large value	123,356	56	56	2	15	11	7	18	4
Mid-cap blend	70,995	29	29	–1	23	9	7	15	3
Mid-cap growth	16,693	19	18	2	26	10	6	16	6
Mid-cap value	35,209	23	23	1	14	10	7	15	5
Small blend	62,292	28	28	0	15	11	9	19	3
Small growth	14,622	17	17	–1	21	11	7	20	5
Small value	16,941	19	19	0	12	12	8	19	6
All equity: Sectors	273,067	243	328	2	42	19	15	32	7
Communications	2,153	14	12	2	47	25	18	40	12
Consumer cyclical	19,222	23	22	–3	39	14	8	25	6
Consumer defensive	16,651	17	17	3	40	13	8	38	6
Energy	22,549	33	32	1	36	13	10	32	5
Financials	36,706	42	41	3	41	25	21	31	7
Health care	40,180	32	30	3	30	11	7	27	7
Industrials	17,210	29	29	–2	43	13	8	33	7
Misc. sector	4,770	24	24	1	51	44	31	28	18
Natural resources	14,146	42	42	4	53	28	22	37	10
Precious metals	7,378	10	10	18	62	39	37	51	6
Real estate	43,117	18	18	2	34	21	17	19	4
Technology	36,051	43	36	1	38	11	7	35	7
Utilities	12,935	16	15	–1	36	10	6	36	5
US bonds: Government	52,556	42	42	4	17	16	14	12	3
Short government	12,650	12	12	0	20	4	3	13	1
Intermediate government	5,319	8	8	6	13	9	6	10	5
Long government	15,116	10	10	7	21	17	15	14	3
Inflation-protected bond	19,471	12	12	6	13	24	22	9	4
US bonds: General	219,342	120	120	20	41	55	50	24	5
Ultrashort bond	9,655	8	8	1	24	13	8	6	2
Short-term bond	40,639	15	15	29	28	45	42	15	3
Intermediate-term bond	68,087	16	16	8	45	32	28	23	3
Long-term bond	2,084	5	5	–5	55	48	42	22	13

(continued)

Table 3. ETF Price Premiums (and Discounts) relative to NAV, January 2007–December 2014 (continued)

Category	Market Cap ($ millions)	No. of Funds: All	No. of Funds: NAV	Avg. Premium (bps)	Volatility of Premium (bps): EW	Volatility of Premium (bps): VW	Volatility of Premium (bps): VW Min.	Bid–Ask Spread (bps): EW	Bid–Ask Spread (bps): VW
Corporate bond	37,710	33	33	31	42	70	66	29	6
High-yield bond	34,494	19	19	31	43	98	92	14	4
Convertibles	2,860	2	2	6	61	70	36	101	13
Preferred stock	16,461	10	10	16	47	63	51	27	8
Bank loan	6,792	4	4	13	24	30	27	8	5
Nontraditional bond	560	8	8	–2	57	38	23	56	38
US bonds: Munis	14,184	33	33	5	64	60	51	23	11
Muni short	4,110	13	13	14	41	37	29	20	10
Muni intermediate	5,828	8	8	19	84	71	62	26	10
Muni long	2,359	9	9	–14	68	50	42	29	12
High-yield muni	1886	3	3	–17	95	116	102	14	12
International equity	379,613	351	345	20	87	84	78	57	6
World stock	17,574	22	22	19	62	49	38	37	11
Foreign large blend	109,051	28	27	27	59	71	67	35	4
Foreign large growth	1,991	6	5	15	59	87	76	39	11
Foreign large value	11,295	24	22	29	80	89	75	38	15
Foreign small/mid-blend	6,820	8	8	37	78	89	70	24	13
Foreign small/mid-growth		1	1	–1	24	24	12	30	30
Foreign small/mid-value	1,104	6	6	22	80	94	81	37	20
Latin America stock	4,976	17	16	18	78	75	71	49	4
Europe stock	32,847	15	14	29	87	62	55	31	9
Diversified Pacific/Asia	2,646	4	4	–4	89	30	23	404	15
Misc. region	29,061	77	77	16	91	105	98	88	9
Japan stock	27,700	19	19	13	122	129	122	31	12
China region	16,683	34	34	22	115	144	133	28	6
India equity	3,804	7	7	17	107	119	112	38	10
Pacific/Asia ex-Japan stock	6,870	13	13	12	92	116	106	42	8
Diversified emerging markets	96,945	59	59	22	81	75	72	70	4
Global real estate	10,245	11	11	25	85	82	71	26	13
International bonds	18,293	44	41	4	75	75	64	47	13
World bond	7,986	25	25	–13	70	48	40	52	16
Emerging market bond	10,307	19	16	31	83	94	81	40	10

(continued)

Table 3. ETF Price Premiums (and Discounts) relative to NAV, January 2007–December 2014 (continued)

Category	Market Cap ($ millions)	No. of Funds: All	No. of Funds: NAV	Avg. Premium (bps)	Volatility of Premium (bps): EW	Volatility of Premium (bps): VW	Volatility of Premium (bps): VW Min.	Bid–Ask Spread (bps): EW	Bid–Ask Spread (bps): VW
Allocation	3,852	43	43	–11	67	42	22	51	26
Conservative allocation	1,204	4	4	–40	99	36	13	45	14
Moderate allocation	1,404	4	4	11	36	17	5	40	15
Aggressive allocation	430	4	4	6	70	66	18	40	22
Target date	116	14	14	–37	92	141	104	59	82
World allocation	642	11	11	18	49	35	23	51	30
Tactical allocation	56	6	6	–7	35	21	10	49	28
Commodities	51,159	45	45	1	98	98	94	43	4
Agriculture	1,197	7	7	–42	127	114	109	104	5
Broad basket	6,014	6	6	17	55	107	101	15	7
Energy	2,492	11	11	0	82	114	59	16	7
Industrial metals	224	3	3	–14	158	229	220	41	16
Precious metals	41,232	18	18	9	102	94	94	45	3
Miscellaneous	51,469	491	377	1	39	35	21	50	8
Currency	2,691	23	23	0	53	45	40	24	7
Long–short	620	11	10	4	102	34	23	40	24
Market neutral	172	11	11	–11	51	34	17	56	33
Multi-alternative	1,658	3	3	5	15	18	7	15	14
Trading (misc.)	1,093	11	11	–4	26	17	6	33	7
Volatility	992	4	4	–1	112	132	135	11	11
Managed futures	213	2	2	24	62	15	9	32	16
Energy limited partnership	9,897	8	8	8	18	8	5	23	8
Leveraged	16,618	81	81	–2	38	47	26	24	9
Bear market	12,815	112	110	–2	33	33	17	27	6
Dead before 2010	5	100	0	0	0	0	0	121	0
Inception in second half of 2014	4,418	110	109	9	32	19	13	26	16
Unmatched	278	15	5	–73	151	91	69	152	32
All	1,968,530	1,813	1,670	6	49	40	36	39	5

Notes: For all ETFs traded in the United States, this table shows the number of ETFs and their last available market cap within each investment category. For the ETFs with available data on NAVs, the table shows the equal-weighted average premium (or discount) of the ETF price (closing midquote) relative to its NAV, as well as the time-series volatility of the premium, either equal weighted (EW) or value weighted (VW) within a category. The "VW Min." column assumes that the market price is any price within the bid and ask so that the distance to the NAV is minimized. The bid–ask spread is the cross-sectional average of the time-series median bid–ask spreads for all ETFs.

To give some idea of transaction costs in the ETFs themselves, the last two columns in Table 3 show their bid–ask spreads. The equal-weighted average is 39 bps and the value-weighted average is only 5 bps, indicating the tremendous trading activity that the larger ETFs have generated. The value-weighted numbers show the lowest bid–ask spreads of 2–5 bps for diversified US equity funds, US government bonds, and commodities, and 5–10 bps for most other categories with at least $1 billion in assets.

The column labeled "VW Min." shows the value-weighted volatility of the premium controlling for the bid–ask spread. I set the price of the ETF between the bid price and the ask price to minimize the absolute value of the premium, effectively assuming that the market is maximally efficient within the bid–ask spread. By construction, the volatility of the premium goes down with such an extreme assumption, but only from 40 bps to 36 bps for the value-weighted results.[10] Hence, transaction costs in ETF shares are unlikely to explain the premiums. This finding should perhaps not be surprising, given that ETFs have been the most liquid and actively traded securities in the equity market in recent years.

How persistent are the premiums? I also computed autocorrelations of the premiums for all categories of funds (unreported). The equal-weighted daily autocorrelation is only 0.30 across all funds, and the average half-life of the premium is 0.58 days. Thus, premiums relative to NAV are short lived, as one might expect when the underlying assets have different trading hours than the ETF. Equity funds have premium half-lives of around 0.5 days, whereas non–Treasury bond funds have half-lives of 2–3 days.

Finally, such premiums as the ones I computed from closing midquote prices can understate the true premiums paid by investors, even when NAVs are not stale.[11] The SEC requires that ETF prospectuses disclose historical premiums computed from closing midquotes, which creates potential incentives for ETF providers to try to influence the closing quotes in order to more closely align the midpoint with the NAV (Gastineau 2009). However, the actual price in the closing auction can differ from the midquote and is not disclosed in ETF prospectuses. Petajisto (2011) reported that premium volatilities computed from closing prices, including only cases where transactions occurred within 15 minutes of market close, are indeed slightly higher than volatilities computed from closing midquote prices: In 2009–2010, the averages were 53 bps versus 49 bps on an equal-weighted basis and 45 bps versus 43 bps on a value-weighted basis. Thus, investors should view the premiums reported in this article as a lower bound; in practice, investors may end up transacting at slightly more volatile premiums.

Premiums and Share Creations. What do historical data suggest about how ETF market makers actually respond to premiums? ETF flows and premiums are both positively autocorrelated, and my hypothesis is that they are linked by the following mechanism: (1) End investors' demand for ETF shares is positively autocorrelated; (2) buy pressure pushes ETF shares to a premium; (3) after the premium is large and persistent enough, authorized participants start conducting arbitrage between ETF shares and the underlying portfolio, creating ETF inflows (increases in shares outstanding); and (4) this arbitrage activity creates sell pressure for ETF shares, which eventually must lead to a reduction in the ETF premium.

Table 4 shows share creations on day *t* as a function of lagged end-of-day premium, with redemptions counted as negative creations. Creations are expressed as a fraction of the average daily trading volume during the same month. Standard errors are computed with double clustering across both funds and time (*t*-statistics based on them are reported in parentheses). This approach takes into account persistent fund-specific effects whereby one fund is trading at a persistent premium (e.g., because of persistent buy pressure combined with illiquid underlying assets), and it also allows premiums to be correlated across similar funds within the same time period.

I found that past premiums positively predict future share creations up to about 10 daily lags (two weeks), with the strongest effect coming from the prior two days. A one-day premium of 1% on a fund would lead to a 4.9% increase in shares

Table 4. Creations and Redemptions as a Function of Lagged ETF Premium, January 2007–December 2014 (*t*-statistics in parentheses)

	(1)	(2)	(3)	(4)	(5)
Premium $t - 1$	0.0493**	0.0325**	0.0217**	0.0207**	0.0216**
	(12.17)	(9.46)	(8.58)	(9.76)	(10.52)
Premium $t - 2$		0.0348**	0.0225**	0.0213**	0.0222**
		(10.16)	(9.77)	(11.21)	(12.17)
Premium $t - 5$ to $t - 3$			0.0131**	0.0116**	0.0123**
			(6.80)	(9.62)	(11.66)
Premium $t - 10$ to $t - 6$				0.0021	0.0032**
				(1.87)	(4.18)
Premium $t - 15$ to $t - 11$					−0.0020*
					(−2.49)
N	1,842,629	1,840,087	1,833,016	1,821,317	1,808,922
R^2	0.4%	0.5%	0.6%	0.7%	0.7%

Notes: The dependent variable is daily ETF shares created or redeemed, expressed as a fraction of the average daily volume of the ETF. The independent variables represent the ETF price premium (%) over the NAV; a premium over multiple days is expressed as the sum of the daily premiums (e.g., the sum of five daily premiums from $t - 15$ to $t - 11$). The *t*-statistics are based on double-clustered standard errors across funds and over time.

*Significant at the 5% level.

**Significant at the 1% level.

outstanding relative to the daily trading volume, and a more persistent one-week premium of 1% would increase shares outstanding by 8% of daily volume. The effect is statistically highly significant. This finding suggests that market makers indeed respond to nonzero premiums within 1–10 days by creating or redeeming ETF shares.

Within style categories, the coefficients are about twice as large for US diversified equity funds, perhaps reflecting their more accurate NAV data. The accuracy of NAV data alone, however, cannot explain differences across categories, because international bond funds have an even larger coefficient and international equity funds are about average. The significance of the results is very similar when scaling share creations by shares outstanding instead of average trading volume or including only funds above $100 million in assets.

Panel A of **Table 5** shows how premiums respond to share creations and redemptions. Creations and redemptions on the same day immediately affect the premium, although by only a small amount: A market maker creating enough new shares to match the daily trading volume (as we saw in Table 2, the median creation is actually 3.1 times daily ETF volume) reduces the premium by about 1 bp by the close of trading, which is statistically significant. Over the following two days, creations continue to reduce the premium by another 1 bp or so; subsequently, they have no statistically significant effect on the premium. This finding suggests that market makers offload their newly created ETF shares in the secondary market immediately before and after the creation process, and thus the price pressure from the new shares arises contemporaneously within about one day of share creation. However, the relatively small size of the effect suggests that shares are created and redeemed for reasons other than arbitrage. Sometimes large investors trade directly in the underlying securities if the expected price impact there is less than in the ETF market; in such cases, the newly created or redeemed shares would not

Table 5. ETF Premium as a Function of Lagged Creations and Redemptions, January 2007–December 2014 (*t*-statistics in parentheses)

	(1)	(2)	(3)	(4)
A. Change in premium from t – 1 to t				
Creations *t*	−0.0103**	−0.0101**	−0.0100**	−0.0099**
	(−6.61)	(−6.58)	(−6.61)	(−6.61)
Creations *t* – 1		−0.0047**	−0.0046**	−0.0044**
		(−3.64)	(−3.61)	(−3.55)
Creations *t* – 2			−0.0026*	−0.0025*
			(−2.16)	(−2.12)
Creations *t* – 3				−0.0021
				(−1.84)
N	1,841,318	1,838,621	1,835,945	1,833,303
R^2	0.0%	0.0%	0.0%	0.0%
B. Level of premium at time t				
Creations *t* – 1	0.0598**	0.0432**	0.0436**	0.0436**
	(20.74)	(18.11)	(18.40)	(18.33)
Creations *t* – 2	0.0563**	0.0403**	0.0403**	0.0404**
	(20.54)	(18.03)	(18.24)	(18.18)
Creations *t* – 3	0.0532**	0.0373**	0.0372**	0.0372**
	(20.53)	(17.63)	(17.87)	(17.91)
Creations: Prior 1 month		0.3046**	0.1457**	0.1590**
		(16.99)	(11.09)	(11.37)
Creations: Prior 3 month			0.0845**	0.0466**
			(12.26)	(6.57)
Creations: Prior 6 month				0.0222**
				(7.25)
N	1,835,178	1,819,207	1,771,856	1,701,115
R^2	0.6%	1.2%	1.4%	1.4%

Notes: The dependent variable in Panel A is the daily change in ETF price premium (%) over the NAV; in Panel B, it is the level of the premium. The independent variables are the ETF shares created or redeemed in the previous three days, expressed as a fraction of the average daily trading volume of the ETF, as well as the cumulative ETF shares created or redeemed in the previous six months, expressed as a fraction of a fund's shares outstanding. The *t*-statistics are based on double-clustered standard errors across funds and over time.

*Significant at the 5% level.

**Significant at the 1% level.

be traded at all in the secondary market and thus would not affect the premium on the ETF.

Panel B of Table 5 shows the long-term relationship between creations and the level of the premium. Creations in the prior three days all significantly predict the level of the premium. In fact, the cumulative creations over the prior one, three, and six months all significantly predict the level of the premium. One explanation for this persistence in creations and premiums is that funds experiencing steady inflows trade at a premium; presumably, investor demand pushes the ETF price to a premium, which then incentivizes market makers to create more ETF shares, but not so aggressively that they would eliminate the premium that is generating their own arbitrage profits. Similarly, the reason an ETF is shrinking is that a market maker is redeeming shares, which is a profitable trade only if it has first purchased those ETF shares in the public market at a discount.

ETF Premiums relative to Peer Groups

In this section, I present my approach to addressing stale NAVs and the results on the remaining ETF mispricings. I also discuss their magnitude in the aggregate and over time, together with recommendations for ETF investors.

Methodology. To resolve the staleness issue in published NAVs, Engle and Sarkar (2006) proposed three econometric models that allowed them to estimate the true NAV. This approach is certainly reasonable, but such models always require assumptions about the price processes involved. In contrast, I used the information in the cross section of ETFs, many of which track identical or nearly identical indexes. Unlike earlier authors, I had the luxury of working with a much bigger cross-sectional sample of funds, making my approach feasible.

I started by sorting ETFs into groups of similar funds, first by grouping funds with the same benchmark index. For all groups with fewer than five funds, I also looked for funds with an economically similar and highly correlated benchmark. This method produced peer groups with anywhere from 2 to 13 funds. I computed each fund's price deviation from its peer-group mean, which I considered the premium on the fund.[12]

For example, assume that a group has five funds that all track the MSCI EAFE index. The funds within the group should move very closely together, regardless of any staleness in their published NAVs. If four of the funds are up 1% and the fifth one is flat, it is likely that the fifth fund is underpriced and will eventually rise back to the same level as its peers.

If these peer-group premiums are zero, that indicates the ETFs are priced efficiently relative to each other and any premiums relative to NAV are due to poor estimates of NAV. Alternatively, if the peer-group premiums are as large as the premiums relative to NAV, I cannot blame NAV staleness and instead conclude that the NAV premiums are all about inefficiency in ETF prices.

As a caveat, I note that this methodology captures any idiosyncratic mispricings on ETFs, but it does not capture a possible systematic mispricing for an entire fund group. It also may add some noise to the premium volatility if funds within a peer group track similar but not identical indexes or if two funds within a group differ slightly in terms of how closely their portfolios replicate the index.

I included inverse ("bear market") and leveraged ETFs with regular ETFs that tracked the same indexes, which required me to delever their returns to obtain comparable return series in which all funds had index betas equal to 1. The delevered fund return R_{del} can be written as a function of the levered fund return R_{lev}, leverage β, and risk-free return R_f:

$$R_{del} = R_f + \frac{R_{lev} - R_f}{\beta}. \quad (1)$$

To reduce the impact of the smallest funds on the results, I eliminated funds below \$10 million in assets as well as funds with daily trading volume less than \$0.1 million, which reduced my sample size from 1,813 to 1,423 funds in 2007–2014.

However, I dropped another 837 funds because I could not find close-enough substitutes for them (funds tracking the same or a highly correlated index) or because they only briefly exceeded the previously mentioned liquidity cutoffs, which reduced the total number of qualifying funds to 586. This sample still covered 87% of all ETF assets; so, from an investment perspective, the qualifying sample can be considered fairly comprehensive.

Estimates of Premiums. **Table 6** shows the volatility of the estimated price premiums on ETFs across the same Morningstar categories as before. To facilitate comparison, premiums are shown side by side with both NAV data and the cross-sectional peer-group method because the sample is slightly different; thus, the NAV premiums are not identical to the ones in Table 3. In other words, for both methods in Table 6, premiums are computed for the same sample size of qualifying funds in peer groups, which also means the average fund in Table 6 is larger than that in Table 3. The equal-weighted volatility of the premium is now 26 bps, which is 30% lower than the 38 bps estimated from NAV data. The value-weighted volatility falls by around 50%, from 37 bps to 18 bps.

Compared with the NAV results, there are a few offsetting effects here as mentioned earlier: The premiums should be smaller because this new method is unaffected by staleness in reported NAVs. At the same time, this method can introduce new noise if the ETF groups contain some funds that are not perfect substitutes for each other in terms of their underlying holdings. Fortunately, one can approximate the magnitude of the noise term by comparing the cross-sectional premium volatilities for US equities with the corresponding NAV premium volatilities because US equities (except perhaps small caps) are not subject to stale-pricing concerns. It turns out that the cross-sectional estimates are about 10 bps higher on an equal-weighted basis but only 2 bps higher on a value-weighted basis, which suggests that the noise resulting from inappropriate group assignment is rather small. Hence, the numbers in Table 6 should be reasonable estimates of idiosyncratic mispricings in ETFs.

The biggest reductions in the volatility of the premium come from illiquid US corporate and long-term muni bonds, international equities and bonds, and precious metals,[13] which are the categories most prone to stale pricing. Nevertheless, there is still nontrivial residual volatility within these ETF groups, with some international equity and bond groups exhibiting volatilities of 50–60 bps, implying 95% pricing bands of over 200 bps. Qualitatively, it is not surprising that the harder and riskier the arbitrage for an authorized participant, the greater the mispricings that remain, but quantitatively the mispricings may still be surprisingly large.

Note that there is no clear relationship between the secondary-market liquidity of the ETF and its premium volatility. For example, ETFs based on high-yield bonds, bank loans, and international equities can be very liquid as measured by low bid–ask spreads (see Table 3) but still have fairly large premium volatilities because the premiums are driven more by the liquidity of the underlying and by other impediments to the creation/redemption arbitrage. Meanwhile, the ETF may be able to manufacture high secondary-market liquidity even from illiquid underlying assets.

The persistence of the premiums goes up when stale NAVs are accounted for. The average autocorrelation rises to 0.70, and the average half-life rises to about two days. The differences across broad categories are surprisingly small, but the economic magnitude matters too: A half-life of two days for a tiny 2 bp mispricing in an S&P 500 fund is not very interesting, but the same half-life for a 50 bp mispricing in a European equity fund *is* interesting. Also within international equity, the "foreign large blend" funds have a premium half-life of only one day, but the Chinese and Indian equity funds have premium half-lives of almost a week.

Economic Magnitude of Mispricings. How large are the dollar amounts involved in these mispricings? Panel A of **Table 7** shows the approximate dollar value of the premiums in actual trades, using both NAV-based premiums and peer-group premiums. I assumed that all trading takes place at the premium (using the bid–ask midpoint) at

Table 6. Cross-Sectional Volatility of Premiums on ETFs, January 2007–December 2014

	Market Cap		No. of Funds		Volatility of Premium (bps)			
Category	All ($ millions)	Groups ($ millions)	All	In Groups	E NAV	E Group	V NAV	V Group
US equity: Diversified	904,995	846,478	301	128	12	22	9	12
Large blend	454,381	445,912	71	33	13	22	9	9
Large growth	110,506	103,306	39	13	12	17	7	9
Large value	123,356	111,660	56	27	10	24	9	20
Mid-cap blend	70,995	64,348	29	13	10	18	9	9
Mid-cap growth	16,693	13,867	19	6	8	15	8	14
Mid-cap value	35,209	16,493	23	8	10	33	9	37
Small blend	62,292	60,393	28	13	16	23	12	14
Small growth	14,622	14,154	17	7	9	18	11	18
Small value	16,941	16,345	19	8	13	26	13	26
All equity: Sectors	273,067	216,895	343	90	18	35	14	29
Communications	2,153	1,354	14	4	57	35	22	26
Consumer cyclical	19,222	11,949	23	4	8	12	10	17
Consumer defensive	16,651	13,073	17	3	9	9	9	10
Energy	22,549	18,130	33	7	6	18	9	21
Financials	36,706	27,812	42	9	18	46	24	42
Health care	40,180	30,077	32	13	13	32	8	26
Industrials	17,210	15,596	29	8	7	33	8	28
Misc. sector	4,770	2,838	24	6	52	76	51	72
Natural resources	14,146	4,980	42	7	15	27	9	23
Precious metals	7,378	7,135	10	4	53	51	39	50
Real estate	43,117	41,276	18	9	14	24	21	22
Technology	36,051	30,886	43	11	8	50	8	39
Utilities	12,935	11,790	16	5	6	21	7	21
US bonds: Government	52,556	45,840	42	25	11	15	15	6
Short government	12,650	9,482	12	4	2	2	2	1
Intermediate government	5,319	5,058	8	7	9	9	8	8
Long government	15,116	14,516	10	7	16	38	16	8
Inflation-protected bond	19,471	16,784	12	7	13	5	23	7
US bonds: General	219,342	183,110	120	50	39	25	52	27
Ultrashort bond	9,655	0	8	0				
Short-term bond	40,639	23,869	15	3	15	11	16	11
Intermediate-term bond	68,087	65,511	16	10	29	15	33	16
Long-term bond	2,084	2,084	5	5	53	26	40	22

(continued)

Table 6. Cross-Sectional Volatility of Premiums on ETFs, January 2007–December 2014 (continued)

	Market Cap		No. of Funds		Volatility of Premium (bps)			
Category	All ($ millions)	Groups ($ millions)	All	In Groups	E NAV	E Group	V NAV	V Group
Corporate bond	37,710	36,398	33	18	38	24	68	34
High-yield bond	34,494	33,137	19	8	48	26	99	36
Convertibles	2,860	0	2	0				
Preferred stock	16,461	15,521	10	3	83	86	68	73
Bank loan	6,792	6,591	4	3	17	12	18	13
Nontraditional bond	560	0	8	0				
US bonds: Munis	14,184	11,225	33	20	72	39	70	39
Muni short	4,110	1,611	13	6	40	24	56	25
Muni intermediate	5,828	5,489	8	4	95	58	74	47
Muni long	2,359	2,320	9	8	74	44	50	38
High-yield muni	1,886	1,804	3	2	113	28	96	28
International equity	379,613	307,861	351	107	83	38	82	24
World stock	17,574	15,365	22	8	49	40	47	39
Foreign large blend	109,051	104,875	28	13	56	26	72	13
Foreign large growth	1,991	0	6	0				
Foreign large value	11,295	5,226	24	10	71	31	88	39
Foreign small/mid-blend	6,820	4,712	8	5	87	50	98	45
Foreign small/mid-growth			1	0				
Foreign small/mid-value	1,104	0	6	0				
Latin America stock	4,976	4,786	17	4	49	24	66	8
Europe stock	32,847	25,373	15	5	76	51	53	28
Diversified Pacific/Asia	2,646	0	4	0				
Misc. region	29,061	5,682	77	18	94	44	113	48
Japan stock	27,700	27,518	19	10	118	33	120	28
China region	16,683	9,767	34	9	115	49	158	51
India equity	3,804	2,668	7	2	117	51	125	56
Pacific/Asia ex-Japan stock	6,870	6,463	13	6	98	44	108	41
Diversified emerging markets	96,945	86,396	59	11	85	35	76	23
Global real estate	10,245	9,028	11	6	57	29	62	31
International bonds	18,293	9,249	44	8	75	42	108	62
World bond	7,986	507	25	2	36	29	36	29
Emerging market bond	10,307	8,741	19	6	88	46	111	63

(continued)

Table 6. Cross-Sectional Volatility of Premiums on ETFs, January 2007–December 2014 (continued)

	Market Cap		No. of Funds		Volatility of Premium (bps)			
Category	All ($ millions)	Groups ($ millions)	All	In Groups	E NAV	E Group	V NAV	V Group
Allocation	3,852	631	43	9	59	36	48	33
Conservative allocation	1,204	553	4	2	18	23	18	23
Moderate allocation	1,404	0	4	0				
Aggressive allocation	430	0	4	0				
Target date	116	78	14	7	71	40	90	46
World allocation	642	0	11	0				
Tactical allocation	56	0	6	0				
Commodities	51,159	41,526	45	9	86	35	92	9
Agriculture	1,197	0	7	0				
Broad basket	6,014	0	6	0				
Energy	2,492	1,946	11	3	44	83	37	75
Industrial metals	224	0	3	0				
Precious metals	41,232	39,580	18	6	107	11	94	5
Miscellaneous	51,469	41,264	491	140	33	14	33	13
Currency	2,691	1,538	23	4	24	9	18	10
Long–short	620	0	11	0				
Market neutral	172	0	11	0				
Multi-alternative	1,658	0	3	0				
Trading (misc.)	1,093	1,053	11	4	10	4	8	3
Volatility	992	458	4	2	130	16	141	16
Managed futures	213	0	2	0				
Energy limited partnership	9,897	9,402	8	3	11	18	8	16
Leveraged	16,618	16,253	81	59	34	16	46	16
Bear market	12,815	12,560	112	67	31	14	33	11
Dead before 2010	5	0	100	1		9		
Inception in second half of 2014	4,418	0	110	0				
Unmatched	278	0	15	0				
All	1,968,530	1,704,080	1,813	586	38	26	37	18

Notes: For all ETFs traded in the United States, this table shows the number of ETFs and their last available market cap within each investment category. From this sample, funds are further assigned to peer groups of 2–13 funds tracking the same or a very similar underlying index. For the funds with a close match that have thus been assigned to groups, the table shows the equal-weighted ("E Group") and value-weighted ("V Group") volatility of the deviation of the fund price from its group mean, averaged across funds within a category. For comparison, the volatility of the NAV premium for the same fund dates is shown in adjacent columns ("E NAV" and "V NAV"). The market price is the bid–ask average at the end of each trading day.

Table 7. Dollar Premiums on ETFs: Trading Volume and Market Capitalization, January 2007–December 2014

Category	No. of Funds		Turnover (%)		Volume ($ millions/day)		Traded Premium ($ millions/year)		
	All	In Groups	All	In Groups	All	In Groups	All NAV	Group NAV	Group
A. Historical ETF premium in actual trades									
US equity: Diversified	296	128	12	12	38,269	38,026	5,284	5,232	3,373
All equity: Sectors	328	90	10	9	11,224	7,891	2,491	2,118	4,871
US bonds: Government	42	25	3	3	1,225	1,162	445	437	128
US bonds: General	120	50	1	1	1,352	1,127	1,117	1,033	456
US bonds: Munis	33	20	1	1	69	53	77	71	33
International equity	345	107	4	4	8,982	7,833	17,274	15,138	2,939
International bonds	41	8	1	1	155	87	170	104	39
Allocation	43	9	2	1	20	3	10	1	1
Commodities	45	9	4	4	2,693	2,525	5,836	5,654	845
Miscellaneous	377	140	26	28	16,038	10,393	8,658	8,454	3,745
Leveraged	81	59	44	45	4,350	4,303	3,217	3,176	1,341
Bear market	110	67	29	30	5,502	5,480	4,085	4,070	2,288
All	1,670	586	8	9	80,026	69,100	41,361	38,243	16,431

Category	No. of Funds		Market Cap ($ millions)		Premium Cap ($ millions)		
	All	In Groups	All	In Groups	All NAV	Group NAV	Group
B. Historical market capitalization of ETF premium							
US equity: Diversified	296	128	441,869	414,093	158	143	246
All equity: Sectors	328	90	143,244	103,435	113	59	194
US bonds: Government	42	25	43,977	39,277	39	37	15
US bonds: General	120	50	126,305	106,390	318	291	139
US bonds: Munis	33	20	8,335	6,322	27	25	12
International equity	345	107	266,515	217,900	1,401	1137	314
International bonds	41	8	12,999	5,929	50	26	10
Allocation	43	9	2,185	406	5	1	1
Commodities	45	9	69,757	58,391	389	350	33
Miscellaneous	377	140	48,238	39,258	94	74	33
Leveraged	81	59	12,046	11,671	31	30	10
Bear market	110	67	17,728	17,283	35	34	12
All	1,670	586	1,163,424	991,402	2,594	2,143	998

Notes: All ETFs traded in the United States ("All") are assigned to peer groups of 2–13 funds tracking the same or a very similar underlying index ("In Groups"). For the funds with a close match that have thus been assigned to groups, Panel A shows the value-weighted average daily turnover of ETF shares, total average daily dollar trading volume, and total absolute dollar premium involved in actual trades each year. The total traded premium is computed in three ways: by using the official price premium relative to NAV for all funds ("All NAV"), the premium relative to NAV only for funds in groups ("Group NAV"), and the peer-group-adjusted price premium for funds in groups ("Group"). Panel B shows the sum of the average market caps across funds and the total absolute average dollar premium; the latter is computed in the same three ways.

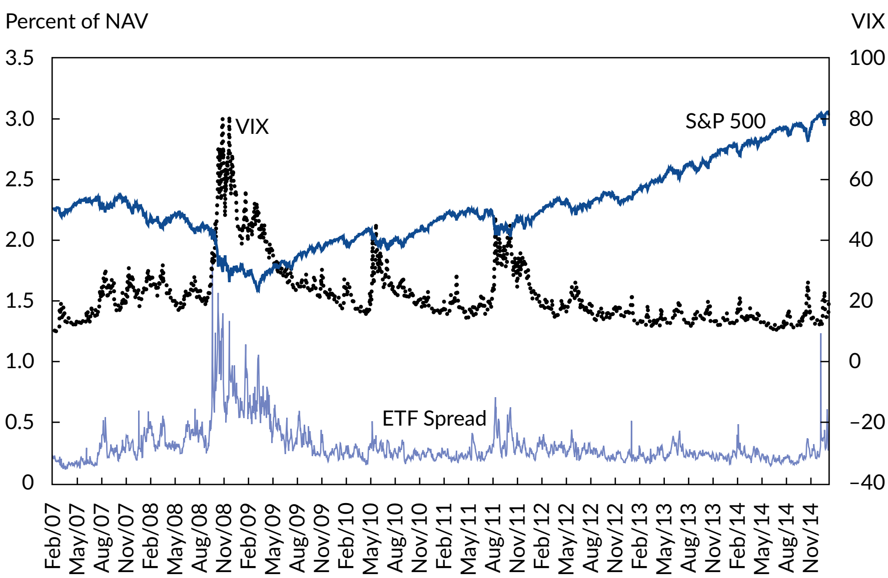
the close, which tends to be more efficient than intraday prices throughout the trading day.

The total premiums across all ETF trades add up to roughly $38 billion a year just for the 586 funds in peer groups, or $41 billion for all funds with NAV data. This result arises partly from a heavy trading volume of $69 billion a day—corresponding to a value-weighted ETF turnover of 9%, or an average holding period of around two weeks—and partly from an average NAV premium of 22 bps per execution. Even the cross-sectional peer-group premiums add up to $16 billion a year, which is economically very significant. In addition, the peer groups miss $11 billion out of the total $80 billion of daily ETF trading volume. If we assume that these excluded funds are mispriced like the funds in the peer groups—which is an unrealistically low estimate because they are less liquid and harder to arbitrage—the traded premiums still add up to almost $20 billion a year across the entire US ETF market. That is approximately how much investors are paying for suboptimal timing of their ETF trades and, conversely, how much investors are earning from liquidity provision.[14] Of course, the same investor may unwittingly end up doing both, but given the large amounts involved, the potential losses from trading with the crowd and the gains from smart liquidity provision highlight the importance of being aware of these issues. For comparison, the entire ETF industry was recently estimated to earn about $6 billion a year in management fees for all US-listed ETFs, which is only one-third of the figure mentioned above.[15]

Panel B of Table 7 shows the average absolute market capitalization of premiums at the close. The average NAV premium within the ETF peer groups is slightly over $2 billion out of an average market capitalization of $991 billion. Adjusting for stale pricing with the peer-group method, the average market value of ETF premiums falls to $1 billion. But because the premiums can fluctuate rapidly, the more relevant metric is the dollar value of the premiums involved in actual trades.

Evolution of Mispricings over Time. One way to measure the efficiency of ETF prices at any point in time is to compute the cross-sectional standard deviation of ETF premiums. This measure is shown in **Figure 3** as the bottom plot, labeled "ETF Spread." In early 2007, the cross-sectional dispersion in premiums starts at about 20–25 bps. It first peaks during the quant crisis of August 2007 but also generally increases afterward, rising to 30–150 bps for most of 2008. After reaching its highest

Figure 3. Cross-Sectional Dispersion of Premium



Notes: This figure shows the cross-sectional standard deviation of the premium across all ETFs at the end of each trading day. The premium is computed relative to a peer-group mean to eliminate any effects from stale pricing. The other plotted time series are the CBOE VIX volatility index and the cumulative return on the S&P 500.

peak in September 2008, it declines again and remains at 30–90 bps for most of 2009. Starting in early 2010, the spread hovers slightly below 30 bps, on average, but there are a number of significant spikes every year of around 100 bps or more. The spikes include such events as the "flash crash" in May 2010, the US debt downgrade in August 2011, the "taper tantrum" of May–June 2013, and a few short-term mini-crashes in 2014. Although the dispersion in premiums is widest during the financial crisis in late 2008, it remains interesting over the last five years of the sample. Going forward, it seems reasonable to assume that the cross-sectional dispersion in premiums will not fall below the present level for at least the next few years.

Why should the dispersion of premiums vary over time? Presumably, it should depend on two things: (1) the trading volume in ETFs as investors move into or out of funds, which generates price pressure for ETF shares, and (2) the amount of arbitrage capital that is able and willing to accommodate that price pressure. Extreme market movements as indicated by the S&P 500 and its volatility might serve as proxies for investors' rebalancing needs. At the same time, the VIX index may serve as a proxy for the availability of arbitrage capital. Figure 3 suggests a link between the dispersion and the VIX index; interestingly, the wide dispersion in ETF premiums preceded the extreme volatility in many key events, such as the quant crisis in August 2007, the Lehman bankruptcy in September 2008, the flash crash in May 2010, and the market jitters in late 2014.

Table 8 measures the relationship between the dispersion in ETF premiums and three different proxies for arbitrage capital: the VIX volatility index, the TED spread, and the average closed-end fund discount. The TED spread is defined as the difference between three-month LIBOR (or Eurodollar) and T-bill rates, which is the premium that a large financial institution would pay for unsecured lending over the true risk-free rate to finance its trading activity (e.g., Brunnermeier, Nagel, and Pedersen 2008). The closed-end fund discount is computed at the end of each trading day as an equal-weighted average discount relative to NAV across all US-listed closed-end funds. It has been used as a measure of investor sentiment (Baker and Wurgler 2006), but it is also plausible that some of the same arbitrageurs operate in both ETF and closed-end fund markets, implying a potentially close relationship between the closed-end fund discount and the ETF premiums and discounts.

Panel A of Table 8 shows that all three measures are related to ETF premiums. Panel B shows that daily changes in each measure similarly explain daily changes in ETF premiums, although with slightly lower statistical significance. The most robust variables are the VIX index and the closed-end fund discount. An increase of 10 percentage points (pps) in the VIX increases the dispersion in ETF premiums by 13 bps, and an increase of 1 pp in the closed-end fund discount increases it by 3 bps in univariate tests. Hence, the funding costs of arbitrageurs and the riskiness of the overall market environment do seem to matter for the efficiency of ETF prices. Furthermore, the efficiency of ETF prices is related to the deviations of closed-end fund values from their NAVs.

Profitability of Active Trading Strategies. If ETFs are never mispriced, any attempt to trade on apparent mispricings will fail to produce a positive alpha even before transaction costs. Thus, the returns to an active trading strategy serve as a convenient summary statistic about the efficiency of market prices. The measurement of the cross-sectional price premium in the previous section naturally lends itself to an active trading strategy: Buy funds trading at a discount relative to their peer group and short funds trading at a premium once the gap becomes sufficiently wide (I assume trading once at the end of each day using the bid–ask average price at the close).

Table 9 shows the portfolio statistics for the trading strategy using data over January 2007–December 2014. The percentage returns are reported for an "unlevered" portfolio that is \$100 long, \$100 short, and \$100 in cash for every \$100 in capital. The excess return on a strategy involving all US-traded ETFs in the peer groups (above the liquidity cutoffs) is 7.00% (t = 13.14), with a very low annual volatility of 1.50% and a Sharpe ratio of 4.68. Controlling for the Carhart model, we see that the strategy is market neutral: It has zero loadings on market, size, value, and

Table 8. Cross-Sectional Dispersion of Premium and Limits of Arbitrage, January 2007–December 2014 (*t*-statistics in parentheses)

	(1)	(2)	(3)	(4)
A. Cross-sectional dispersion of premium				
VIX index	0.0133**			0.0100**
	(16.83)			(13.79)
TED spread		0.1917**		0.0487**
		(9.93)		(3.18)
CEF discount			0.0322**	0.0100**
			(9.37)	(6.61)
N	1,975	1,967	1,975	1,967
B. Daily change in cross-sectional dispersion of premium				
Δ(VIX index)	0.0097**			0.0065**
	(13.64)			(8.17)
Δ(TED spread)		0.1496**		0.0837**
		(6.77)		(3.88)
Δ(CEF discount)			0.0381**	0.0244**
			(13.78)	(6.98)
N	1,974	1,958	1,974	1,958

Notes: The dependent variable in the regressions in Panel A is the cross-sectional standard deviation of the premium across all large ETFs at the end of each trading day. Large funds are defined as having at least $100 million in assets. The premium is computed relative to a peer-group mean to eliminate any effects from stale pricing. The explanatory variables are the CBOE VIX volatility index, the spread between three-month LIBOR and T-bill rates, and the average equal-weighted discount (relative to NAV) on all US-traded closed-end funds. Panel B shows similar regressions, except that both the dependent and the independent variables are expressed as changes from the previous day. The *t*-statistics are based on Newey–West standard errors with five lags.

**Significant at the 1% level.

momentum, with a Carhart alpha of 6.84% (t = 13.60) a year and an information ratio of 4.81.

Investigating the trading more closely, I found that the profits tend to come from international funds and illiquid underlying securities, consistent with the results in Table 6, whereas diversified US equities, US Treasury bonds, short-term bonds, and commodities (mostly gold and silver) tend to produce very modest returns. Sector funds are somewhere in the middle, with some sectors priced more efficiently than others. Removing diversified US equities, Treasury bonds, and commodities,[16] the Carhart alpha rises to 10.33% (t = 14.44) a year, although volatility also rises slightly, whereas the information ratio inches up to 5.10. Excluding also the sector funds, the Carhart alpha rises to 16.19% (t = 13.90) a year, again with a similar information ratio of 4.86. Economically, this result means that my simple rule identifies mispriced ETFs that will converge to their fundamental values at a rate of 6.4 bps a day. Because most positions are held for longer than a day, the level of mispricing can rise to a multiple of that amount.

These returns to active strategies seem attractive. However, a real-life implementation of the strategy could add a few complications. First,

Table 9. Profitability of Trading against ETF Mispricings, January 2007–December 2014 (*t*-statistics in parentheses)

				Beta					
Benchmark Model	Intercept (annual %)	Information Ratio (annual)	Residual Volatility (annual %)	Mkt. Rf.	SMB	HML	UMD	df	R^2
A. Including all funds									
None	7.00	4.68	1.50					1973	0.0%
	(13.14)								
CAPM	6.84	4.76	1.44	0.02				1972	7.9
	(13.58)			(4.90)					
FF	6.84	4.81	1.42	0.02	−0.02	−0.01		1970	9.9
	(13.62)			(4.80)	(−2.33)	(−1.04)			
Carhart	6.84	4.81	1.42	0.02	−0.02	−0.01	0.00	1969	9.9
	(13.60)			(4.92)	(−2.33)	(−0.90)	(0.31)		
B. Excluding diversified US equities, US government bonds, commodities, miscellaneous									
None	10.50	5.00	2.10					1973	0.0
	(14.05)								
CAPM	10.32	5.06	2.04	0.02				1972	5.9
	(14.43)			(4.50)					
FF	10.33	5.10	2.03	0.03	−0.03	−0.01		1970	7.2
	(14.46)			(3.91)	(−2.08)	(−0.39)			
Carhart	10.33	5.10	2.03	0.03	−0.03	−0.01	0.00	1969	7.3
	(14.44)			(3.90)	(−2.06)	(−0.51)	(−0.56)		
C. Excluding diversified US equities, US government bonds, sector funds, commodities, miscellaneous									
None	16.41	4.82	3.40					1973	0.0
	(13.55)								
CAPM	16.19	4.83	3.35	0.03				1972	3.2
	(13.86)			(3.05)					
FF	16.18	4.86	3.33	0.03	−0.04	−0.02		1970	4.4
	(13.91)			(2.73)	(−1.85)	(−0.72)			
Carhart	16.19	4.86	3.33	0.0	−0.04	−0.02	−0.01	1969	4.5
	(13.90)			(2.70)	(−1.84)	(−0.85)	(−0.74)		

Notes: This table shows the returns on a fully invested but unlevered long–short portfolio that takes positions against the estimated mispricings. The *t*-statistics are based on White's standard errors. For these calculations, levered and inverse funds are counted as part of their underlying style, so they are included with other funds that track the same index (and thus are not excluded with the "Miscellaneous" category).

there may not be enough trading volume in some ETFs to make the strategy interesting. Second, even when trading volume is sufficiently high, it may occur at different times during the day for different funds, and this nonsynchronicity may introduce the false appearance of profitability. Third, the profits are sensitive to transaction costs, so the execution strategy plays a key role.

To address these concerns, I repeated the calculations with an intraday dataset using five-minute periods from 9:30 a.m. to 4:00 p.m. I constructed a real-time signal on the basis of currently observable prices and then traded subsequently on that basis, fully addressing potential issues of nonsynchronous trading. I also recomputed trading profits assuming trading at actual transaction prices (five-minute volume-weighted average price) and with a maximum participation rate of 10%. This participation rate constraint on maximum trading volume implies that larger portfolios will be less profitable because there will not be enough volume to allow us to reach our ideal target position in some ETFs. I found that the strategy remains very profitable with intraday trading at actual transaction prices, but the capacity is somewhat limited. For example, the annual information ratio for a $100 million long–short portfolio falls to about 2. Furthermore, the strategy should not be executed aggressively because paying the full bid–ask spread (buying at the ask, selling at the bid) each time would significantly reduce its profitability; instead, it should be run as a passive market-making strategy with constantly updated limit orders, which is feasible because it uses a broad cross section of hundreds of ETFs, potentially trading in any of them at any point in time. In fact, being a liquidity provider could even enhance the profits of the strategy up to a certain dollar capacity.

Regardless of one's view on the exact level of the information ratio after implementation costs, an important implication for market efficiency remains: These trading profits document that the actual prices faced by ETF investors can differ significantly from the true value of the underlying portfolio, thus presenting a potentially large hidden cost for ETF investors.

How Investors Should Trade ETFs. The evidence in this article suggests that trading US-listed ETFs on diversified US equity indexes, nominal US Treasury bonds, or precious metals is easy and usually harmless for the average investor, with the exception of occasional crisis periods. However, trading ETFs with non-Treasury bonds or international securities as the underlying assets exposes the investor to the risk of poor trade timing because of premiums. One approach is to compare the ETF's price with the intraday indicative value (IIV), which is an intraday estimate of the ETF's NAV based on the latest prices of the ETF creation basket and published by the exchange every 15 seconds. But when the underlying securities have not traded because of illiquidity or time zone differences, the IIV can also be a stale measure of portfolio value, which tends to be a problem precisely for the funds that are most prone to mispricing. Thus, some professional investors have developed their own proprietary IIV estimates (e.g., using futures prices). Alternatively, the peer-group approach presented in this article is a reasonable way to check whether a particular ETF has become cheap or expensive relative to its peers in recent days, although even this approach is not conveniently accessible for most retail investors. Another, simpler approach is to look at the latest official premiums but trade only when markets have been flat for the last few days, because then even stale NAVs have had a chance to catch up with the latest market prices of the underlying portfolio.

Conclusion

The dramatic growth of the ETF market since 2006 has brought these investment vehicles to a large fraction of relatively unsophisticated individual investors. It is easy for an investor to fall into the trap of focusing so much on the expense ratios of funds that the transaction price for ETF shares is overlooked. Given that US ETF assets were about $2 trillion and growing in 2014, any nontrivial mispricing in ETFs has the potential to represent a considerable wealth transfer from less sophisticated individual investors to more sophisticated institutional investors.

In this article, I have provided new empirical evidence on the state of market efficiency in ETFs. Funds holding liquid domestic securities are priced relatively efficiently, whereas funds with international or illiquid holdings exhibit nontrivial premiums relative to NAVs, which is qualitatively consistent with the costs and uncertainty faced by arbitrageurs in these funds. More surprisingly, US sector funds holding liquid domestic stocks can also exhibit nontrivial premiums.

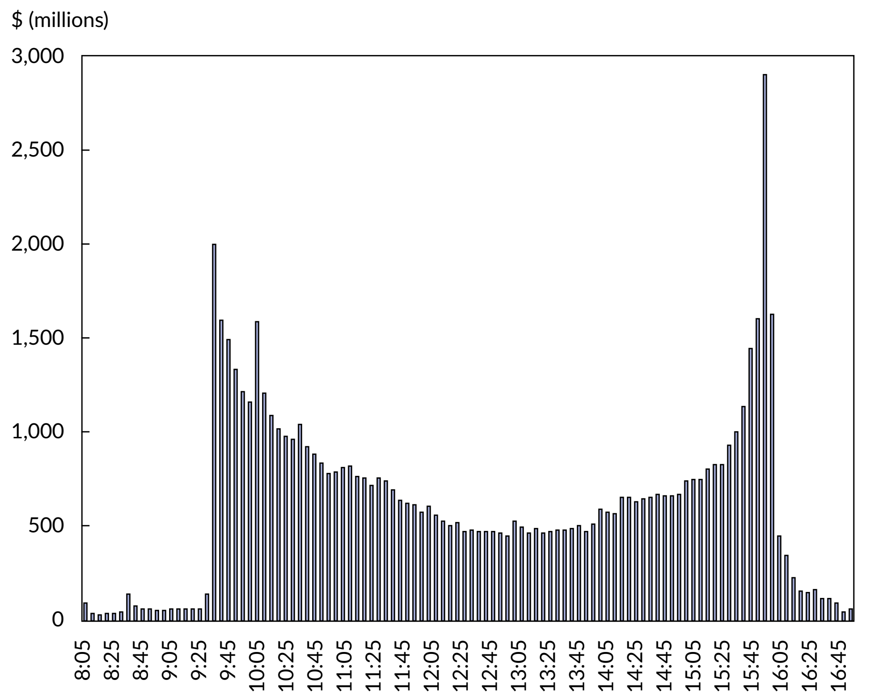
I have proposed a new approach to detecting mispricings in ETFs: Instead of comparing ETF prices with NAVs, measure them relative to the current market prices of a peer group of similar funds. This approach eliminates the problem of stale NAVs. I found that this adjustment reduces the premiums on funds with international or illiquid holdings but still leaves them fluctuating within a pricing band of 100–200 bps, which is economically significant, indicating that nontrivial mispricings remain. This result is confirmed by tests involving the creation of an active trading strategy to exploit these mispricings, because the strategy produces economically significant profits before transaction costs with a high degree of statistical significance as well.

ETFs are convenient vehicles for accessing various market segments and generally come with many virtues, such as low expense ratios, low turnover (implying low transaction costs paid by the fund), and high tax efficiency, so they have legitimately earned their place in the market. However, cost-conscious individual investors should be aware of the potential to transact at a disadvantageous price and how to avoid it so they can fully capture the benefits of these new investment vehicles.

Appendix A. Intraday Trading

Figure A1 shows the total trading volume in all ETFs in five-minute periods throughout the day, averaged across all trading days in 2010, which is the last complete year in my intraday sample. ETFs exhibit the same type of clustering as other securities: Most of the volume occurs at the beginning and end of the trading day. In the middle of the day, trading intensity is about 30%–50% of the value near the beginning and end of the day, but it is certainly still at an economically meaningful level. Anecdotal observations suggest that some ETFs tend to search for their efficient prices

Figure A1. Daily Trading Volume in 2010



Note: This figure shows the total trading volume ($ millions) across all US-listed ETFs for five-minute intervals during the day, averaged across all trading days in 2010.

early in the trading day and then become more efficiently priced toward the close, but that does not seem to hinder overall trading activity in the morning, when trading is essentially just as intense as at the end of the day.

Before November 2008, many ETFs, including SPY, were trading on AMEX until 4:15 p.m. ET, or 15 minutes longer than the equity securities on which the index values were based. Since then, exchange trading hours for ETFs have become standardized, starting at 9:30 a.m. and ending at 4:00 p.m., which presumably was driven by the NYSE's acquisition of AMEX. Fixed-income ETFs still close at a different time than the underlying securities (4:00 p.m. rather than 3:00 p.m. for the bond market), and funds based on international securities will always close at different times than their underlying securities.

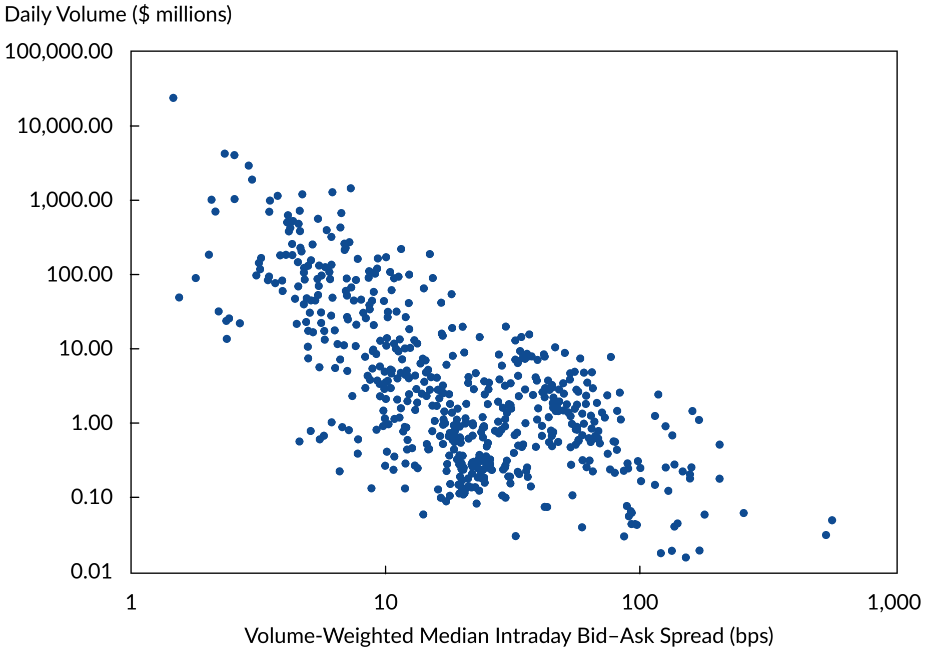
How liquid are ETFs in general? **Figure A2** shows the average daily volume for all ETFs, plotted against their volume-weighted median intraday bid–ask spreads. To capture the liquidity that a typical investor would face, I looked at intraday spreads and not closing spreads,[17] computing the volume-weighted median for each fund to reflect the spreads at the time that actual trades were occurring. I found that all funds with bid–ask spreads below 10 bps also have at least $10 million in daily trading volume; conversely, the dozen funds with over $1 billion in daily trading volume all have spreads at or below 10 bps. More surprisingly, among the funds with a median spread of about 100 bps or above, there are still several funds with over $1 million in trading volume; this spread seems rather large, given that these investment vehicles are fully transparent and passively managed. For the average investor, these findings also highlight the importance of efficient trade execution, especially for short average holding periods.

Editor's Note

Submitted 25 September 2013

Accepted 6 October 2016 by Robert Litterman

Figure A2. Trading Volume and Median Intraday Bid–Ask Spread



Notes: This figure shows the mean daily trading volume plotted against the volume-weighted median bid–ask spread for all US-listed ETFs in 2010. The numbers are based on intraday five-minute intervals from 9:30 a.m. to 4:00 p.m. Both axes are in log scale.

Notes

1. A more recent headline-grabbing episode occurred on 24 August 2015, when even very large funds like IVV, with $70 billion in assets, were briefly trading at around a 20% discount to NAV.

2. For purposes of illustration, I am assuming in these first few paragraphs that the NAV accurately reflects the market value of the underlying portfolio. However, the mutual fund literature makes clear that this is not the case with some types of funds (e.g., Chalmers, Edelen, and Kadlec 2001; Goetzmann, Ivkovic, and Rouwenhorst 2001; Zitzewitz 2003).

3. Worldwide in December 2014, the number of live ETFs, including exchange-traded notes and exchange-traded commodity funds, reached 5,428, with an estimated $2.8 trillion in assets (BlackRock 2014).

4. In a more recent study, Madhavan and Sobczyk (2014) adjusted for stale pricing by using a state-space model somewhat similar to that of Engle and Sarkar (2006), but they focused on quantifying the magnitude and speed of ETF price discovery across the four broad categories of domestic versus international and equity versus fixed-income funds. In addition, a few authors have analyzed other separate but related questions, such as whether style investing creates co-movement for premiums in US equity ETFs (Broman 2016) and whether premiums in equity ETFs are corrected overnight or intraday (Fulkerson and Jordan 2013).

5. A few highest-fee funds charge fixed-dollar expenses, which can result in enormous percentage fees even though they have only a very small amount of assets.

6. These fees represent the maximum possible cost for an authorized participant. If the actual transaction costs are lower, the authorized participant will typically have to pay only the smaller amount. If transaction costs are expected to be higher than these maximum amounts (e.g., during a temporary lack of liquidity in the underlying market), the fund sponsor may refuse cash creations altogether.

7. In CRSP, I found 5 data points (out of about 600,000) in which the daily ETF price was off by a nontrivial amount, and OpenTick had about 20 such data points. I could not find similar errors in the iShares data. I could not set the cutoff much lower than 20% because there were several legitimate data points in which the premium was greater than 10%.

8. There were 1,435 live funds and 378 dead funds as of December 2014. Of the dead funds, 100 died before my first Morningstar data snapshot in March 2010, so they are in the "dead before 2010" category; the remaining 278 funds that died later during 2010–2014 have category data and are thus in the appropriate categories among live funds. The NAV data cover essentially all live funds (1,432 out of 1,435) and a majority of dead funds (238 out of 378).

9. Piccotti (2015) suggested that the positive average premiums for international funds and bond funds might reflect the cost of access that investors are rationally willing to pay in partly segmented markets. In general, the premium distribution is fairly symmetrical, although it has fat tails (Petajisto 2011).

10. Because the value-weighted spreads are tighter than the equal-weighted spreads, the value-weighted results are more informative about potentially inefficient pricing in this case, where we allow the "true" price to be anywhere within the spread.

11. I thank an anonymous referee for pointing that out.

12. The largest group comprised the funds (both levered and unlevered) tracking the S&P 500. My peer groups were usually, but not always, within a single Morningstar style category. I computed the total return index of each fund and normalized the index to start at the same value for all funds within a group. I computed price deviations relative to their 30-day moving average so that any long-term differences in how two funds tracked an index (e.g., because of differences in costs or underlying ETF portfolios) did not meaningfully affect the premium estimates.

13. Precious metals suffer from stale pricing because NAVs are based on spot prices of gold and silver that are determined at 12:00 p.m. or 3:00 p.m. local time in London, which is six to nine hours before the ETF market close at 4:00 p.m. in New York.

14. Furthermore, an investor demanding liquidity may end up paying the bid–ask spread (i.e., buying at the ask and selling at the bid), which is another cost in addition to the ETF midquote being mispriced. Conversely, a patient liquidity provider may be able to earn part of the spread and thus benefit even more.

15. See Dave Nadig, "Here's How Much Money the ETF Industry Makes," ETF.com (6 October 2016).

16. I also dropped the "Allocation" and "Miscellaneous" groups to keep things simple. However, levered and inverse funds in this case are grouped with their underlying asset class, not as part of the "Miscellaneous" category, so they are included in this trading strategy.

17. Closing spreads would be about 20% lower than intraday spreads, on average.

References

Ackert, Lucy F., and Yisong S. Tian. 2000. "Arbitrage and Valuation in the Market for Standard and Poor's Depositary Receipts." *Financial Management*, vol. 29: 71–88.

———. 2008. "Arbitrage, Liquidity, and the Valuation of Exchange-Traded Funds." *Financial Markets, Institutions & Instruments*, vol. 17, no. 5 (December): 331–362.

Baker, Malcolm, and Jeffrey Wurgler. 2006. "Investor Sentiment and the Cross-Section of Stock Returns." *Journal of Finance*, vol. 61: 1645–1680.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi. 2014. "Do ETFs Increase Volatility?" Working paper, Ohio State University.

BlackRock. 2014. "ETP Landscape" (December).

Broman, Markus S. 2016. "Liquidity, Style Investing, and Excess Comovement of Exchange-Traded Fund Returns." *Journal of Financial Markets*, vol. 30 (September): 27–53.

Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen. 2008. "Carry Trades and Currency Crashes." *NBER Macroeconomics Annual*, vol. 23, no. 1: 313–347.

Chalmers, John M.R., Roger M. Edelen, and Gregory B. Kadlec. 2001. "On the Perils of Financial Intermediaries Setting Security Prices: The Mutual Fund Wild Card Option." *Journal of Finance*, vol. 56: 2209–2236.

Delcoure, Natalya, and Maosen Zhong. 2007. "On the Premiums of iShares." *Journal of Empirical Finance*, vol. 14: 168–195.

Elton, Edwin J., Martin J. Gruber, George Comer, and Kai Li. 2002. "Spiders: Where Are the Bugs?" *Journal of Business*, vol. 75: 453–472.

Engle, Robert, and Debojyoti Sarkar. 2006. "Premiums-Discounts and Exchange-Traded Funds." *Journal of Derivatives*, vol. 13, no. 4: 27–45.

Fulkerson, Jon A., and Bradford D. Jordan. 2013. "Reading Tomorrow's Newspaper: Predictability in ETF Returns." *Journal of Index Investing*, vol. 4, no. 1 (Summer): 23–31.

Gastineau, Gary. 2009. "How to Minimize Your Cost of Trading ETFs." *Journal of Indexes*, vol. 12: 24–31.

———. 2010. *The Exchange-Traded Funds Manual*, 2nd ed. Hoboken, NJ: John Wiley & Sons.

Goetzmann, William N., Zoran Ivkovic, and K. Geert Rouwenhorst. 2001. "Day Trading International Mutual Funds: Evidence and Policy Solutions." *Journal of Financial and Quantitative Analysis*, vol. 36: 287–309.

Madhavan, Ananth, and Aleksander Sobczyk. 2014. "Price Dynamics and Liquidity of Exchange-Traded Funds." Working paper, BlackRock.

Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti. 2010. "ETF Arbitrage." Working paper, Massey University (November).

Petajisto, Antti. 2011. "Characteristics of Market Trading in Exchange-Traded Funds: 2009–2010." Research report (August): http://www.petajisto.net/papers/etfev15.pdf.

Piccotti, Louis R. 2015. "ETF Premiums and Liquidity Segmentation." Working paper, SUNY Albany (7 November).

Poterba, James M., and John B. Shoven. 2002. "Exchange-Traded Funds: A New Investment Option for Taxable Investors." *American Economic Review*, vol. 92, no. 2 (May): 422–427.

Sullivan, Rodney N., and James X. Xiong. 2012. "How Index Trading Increases Market Vulnerability." *Financial Analysts Journal*, vol. 68, no. 2 (March/April): 70–84.

Tucker, Matthew, and Stephen Laipply. 2010. "Understanding Bond ETF Premiums and Discounts: A Conceptual Framework." *Journal of Indexes* (September/October): 40–48.

Wurgler, Jeffrey. 2011. "On the Economic Consequences of Index-Linked Investing." In *Challenges to Business in the Twenty-First Century*. Edited by Gerald Rosenfeld, Jay Lorsch, and Rakesh Khurana. Cambridge, MA: American Academy of Arts and Sciences.

Zitzewitz, Eric. 2003. "Who Cares about Shareholders? Arbitrage-Proofing Mutual Funds." *Journal of Law Economics and Organization*, vol. 19, no. 2 (October): 245–280.