When the bellwether dances to noise: Evidence from exchange-traded funds

Zhi Da and Sophie Shive*

September 2, 2013

Abstract

We provide novel evidence that arbitrageurs can exacerbate return comovement via ETF arbitrage. Using a large sample of U.S. equity ETF holdings, we find a strong relation between measures of ETF activity and return comovement at both the fund and the stock levels, after controlling for a host of variables and fixed effects. The effect is stronger among small and illiquid stocks and during market turbulence. An examination of delay measures and mutual-fund-flow-induced price pressure suggests that at least some ETF-driven return comovement is excessive. In other words, ETFs may reduce diversification, the very benefit they were designed to facilitate.

*Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556. Tel. 574-631-8301. Fax. 574-631-5544. Email zda@nd.edu or sshive1@nd.edu. Thanks to participants in the State of Indiana Finance Conference, 2013 China International Conference in Finance, and seminar at University of Notre Dame, Malcom Baker, Robert Battalio, Martijn Cremers, Ben Golez, Robin Greenwood, Bing Han, Paul Schultz, David Solomon, Mao Ye, and Xiaoyan Zhang. This article has been previously circulated under the title "Exchange Traded Funds and Asset Return Correlations." Errors are ours.

1 Introduction

Perhaps due to a half-century of encouragement from finance academics, investment assets are increasingly indexed, but the implications for asset prices of large amounts of indexed investment are not well understood. Citing evidence of mispricing and increased correlations among asset returns, Wurgler (2010) warns that over-indexing may result in contagion and mispricing risk. Exchange-traded funds (ETFs), baskets of equities traded on an exchange like stocks, are a growing asset class that has made indexing cheaper and more convenient for many investors. U.S.-based exchange-traded funds had $1.2 trillion in assets under management by the end of 2012,[1] and net cash flows into equity exchange traded funds have been $71, $64, and $89 billion in 2010, 2011 and 2012, while equity mutual fund flows were -$23, -$128, and -$122 billion.[2] Since these funds will by all measures play a large role in the future of saving and investing, it is important to understand if and how they will affect prices, both in absolute and compared to traditional mutual funds and institutions.

Along with information, ETFs have a potential to transmit non-fundamental shocks - be it liquidity shocks or sentiment-related mispricing - via the arbitrage that they make possible between the ETF and the underlying basket of shares. If demand curves for stocks are downward-sloping, arbitrageurs can move stock prices by simultaneously taking opposite positions in the ETF and the underlying basket of stocks. This arbitrage trading can lead to two adverse effects. First, stocks held by ETFs might comove more with each other than their fundamentals would suggest. Second, due to their relative convenience, ETFs could be prone to sentiment shocks and other mispricing risk, and pass this on to their component stocks. Arbitrageurs, who are generally enforcers of price efficiency, can thus at times make "mispricing" more systematic, consistent with results in Shleifer and Vishny (1997), Hong, Kubik, and Fishman (2012) and Lou and Polk (2013). Ironically, ETFs may reduce diversification, the very benefit they were designed to facilitate.

[1] http://www.etftrends.com/2012/03/etf-assets-up-115-billion-this-year/

[2] Source: Deutsche Bank, Bloomberg Finance LP 2012 figures are as of end of November 2012. http://www.etf.db.com/DE/pdf/DE/research/researchglobal_2013_01_11.pdf

Using a large cross-section of 699 US equity ETFs and almost 4,700 stocks from July 2006 to June 2012, we find strong evidence that ETFs contribute to equity return comovement and mispricing risk. An ETF-level analysis reveals that the higher an ETF's turnover, the more its component stocks move together, controlling for fund and time fixed effects in addition to a host of fund-level control variables. Fund fixed effects alleviate the selection bias where similar stocks are selected by the same ETF. Time fixed effects are also crucial since both ETF activities and stock comovement can be driven by the same macroeconomic variables.[3] Our analysis also corrects for cross-correlation in error terms arising from common holdings across ETFs.

At the fund level, a one-standard-deviation increase in the turnover of a typical ETF in our sample increases the average correlation among its component stocks by 1.2%. This effect is stronger among larger ETFs and ETFs with turnover that closely tracks the turnover of its underlying stocks, supporting the arbitrage trading mechanism. Limits-to-arbitrage alleviate this effect. Corporate bonds and municipal bonds are more costly to trade than equity, and we do not find a link between ETF turnover and return correlation among ETFs that hold significant amounts of such bonds.[4]

We also investigate whether ETF creation and redemption activity affects stock comovement. ETFs are created and redeemed in kind, and this activity might induce correlated trading in the portfolio. We find that our measure of creation and redemption activity is less strongly related to comovement than are ownership or turnover. This is not surprising. As explained in Abner (2010), creation and redemption is more onerous than trading and typically occurs after market close, when an arbitrage opportunity may no longer be available.

We also conduct our analysis at the stock level. First, we find that the higher the total ETF ownership of a stock, the more it comoves with the market in the subsequent month. This holds controlling for stock and time fixed effects and a host of stock-level control variables. For example,

[3] For example, Forbes and Rigobon (2002) show that equity correlation tends to increase during volatile periods when the trading volumes are also high.

[4] We do not use the daily difference between ETF price and ETF NAV as a proxy for arbitrage trading for two reasons. First, there is potential non-synchronicity issue between the ETF price and its NAV, making their difference a noisy measure of mispricing. Second and more importantly, as is shown by Gagnon and Karolyi (2010), a price difference can reflect either an actual opportunity for arbitrage trade or the presence of limits-to-arbitrage.

a 1% increase in ETF ownership of a stock is associated with an 0.03 increase in CAPM beta. Importantly, the effect of ETF holdings is more than three times larger than the effect of mutual fund holdings or other institutional holdings of the stock and is not driven by flow-induced mutual fund trading as in Lou (2012). Second, as in the fund-level analysis, we also find that the weighted average turnover of the ETFs that own the stock is related to how much the stock comoves with the market. A one-standard-deviation increase in weighted average turnover is associated with an 0.081 increase in a stock's CAPM beta, again controlling for stock and time fixed effects and a host of stock-level control variables. Finally, the effect of ETF activities on stock comovement is stronger among small stocks, stocks with low turnover, and during turbulent times when mispricings are ample.

Given the evidence for a positive link between ETF activities and return comovement, a natural question follows: does the increased return comovement reflect faster incorporation of common information in the market that ETF trading helps to facilitate; or does it reflect "excessive" price movement due to non-fundamental shocks that ETF trading helps to propagate?

We examine this important question at the fund and stock levels. At the fund level, we use a relatively exogenous non-fundamental shock that has been widely studied in the literature, price pressure coming from mutual fund flows.[5] We regress the correlation of daily returns of the two largest fund holdings on the absolute value of the flow-induced price pressure of the remaining holdings (i.e., stocks connected via the common ETF ownership). We find that a large unidirectional push from mutual fund flows on the connected stocks causes the largest two stocks in the fund to comove more, consistent with the resulting arbitrage activity that is encouraged by this push.

Next, we perform a stock-level test by looking at lagged market betas. Empirically, we find that stocks with higher measures of ETF activity tend to have significantly negative betas on lagged market returns, and that a stock's lagged betas on market returns are negatively related to the activity of ETFs owning the stock. This suggests that ETF activity is related to overshooting and

[5]For example, Edmans, Goldstein, and Jiang (2012) as well as Khan, Kogan, and Serafeim (2012) use the price pressure from mutual fund trading as an instrument to study corporate activities such as SEO and acquisition.

reversals in prices, a symptom of "excess" comovement. In sharp contrast, under the alternative incorporation of common information hypothesis, the lagged betas should never be negative.

Our paper is related to the large literature on return comovement, spanning the credit, the commodity, and the equity asset classes. In equities, much work has found that adding a stock to an index affects its price (Harris and Gurel (1986), Shleifer (1986), Lynch and Mendenhall (1997), Wurgler and Zhuravskaya (2002), Kaul, Mehrotra, and Morck (2002)) and correlation between the newly added stocks and the stocks in the index increases (Goetzmann and Massa (2003) and Barberis, Shleifer, and Wurgler (2005) for the S&P 500 and Greenwood and Sosner (2007) for the Nikkei 225). This literature is subject to the potential caveat that missing fundamental factors are driving both the index addition and deletion decision and comovement.[6] Examining arbitrage-driven ETF turnover helps to alleviate this concern. In addition, Barberis and Shleifer (2003) and Peng and Xiong (2006) argue that categorical learning and investing by investors could lead to excessive comovement among stocks with similar characteristics or styles. ETFs, by making it easier to trade stocks with similar characteristics, could potentially contribute to style-based return comovements. Finally, a recent literature has linked correlated institutional ownership and trading to excessive return comovement. Among others, it includes Greenwood and Thesmar (2011), Anton and Polk (2013) and Bartram, Griffin, Lim, and Ng (2013) in the international context. To the extent that institutions have some discretion in deciding when and what to trade, ETF arbitrage is more likely to drive return comovement among its component stocks.

Our paper is also related to the growing literature on ETFs. Boehmer and Boehmer (2003) find that the initiation of trading of three ETFs on the NYSE increased liquidity and market quality. Hamm (2011) finds a positive relation between ETF ownership percentage and a stock's liquidity, especially for stocks held by highly diversified ETFs. Engle and Sarkar (2002), Petajisto (2011), and Marshall, Nguyen, and Visaltanachoti (2012) focus on the drivers of differences between the market price of the ETF and the price of the underlying portfolio, and Jiang and Yan (2012)

[6] Greenwood (2008) that takes advantage of the index weighting scheme is a notable exception.

investigate levered ETFs. Our paper extends this stream of literature by examining the impact of ETF-underlying arbitrage on return comovement and mispricing risk.

A recent study by Ben-David, Franzoni, and Moussawi (2012) provides interesting examples where arbitrage activity propagates liquidity shocks from ETFs to the underlying stocks and increases volatility, but does not investigate stock comovement and mispricing risk. Therefore, its findings do not directly speak to the impact of ETFs on diversification which is the focus of our paper. In addition, we examine different aspects of ETF activities (holdings, creation/ redemption, and trading) consistently in the same model. In another contemporaneous study using proprietary daily holdings data on 12 ETFs, Staer (2012) confirms the positive relation between ETF turnover and return comovement at higher frequency. In contrast to his tests, our study covers a much broader cross-section including 699 ETFs and 4,700 stocks. The broader coverage allows us to conduct tests at both the fund level and the stock level. It also improves the precision associated with the stock-level tests. For example, as of 2011, Apple Inc. appears in the top 10 holdings in more than 57 ETFs, IBM in 52 ETFs, and WalMart in 30 ETFs.[7] To measure the aggregate exposure of a stock to ETF activities, it is therefore important to account for a broad cross-section of ETFs.

Finally, our work also has important policy implications. The ETF structure has many benefits. Among others, ETFs provide a cheaper and more efficient way for investors to invest in a broad market index. However, with the growth and popularity of ETFs, regulators have become concerned about their potential distorting effect on markets. For example, Bradley and Litan (2011) in their testimony before the Senate Banking Committee, argued: "High co-movement of securities is not new... What is new, however, is how ETFs decrease diversification benefits, with stocks and sectors worldwide moving together, even when there is no panic (Page 6)." Sullivan and Xiong (2012) also note that the recent rise in the importance of ETFs and other index trading activities coincides with rising return correlations among stocks. Taking advantage of a large panel of ETF holdings, we

[7]see www.etfdb.com.

are able to better quantify the incremental impact of ETFs on equity comovement and mispricing risk. Thus, our findings are clearly relevant for the developing ETF market. While the effects at current levels of ETF holdings (2.4% of a stock's market capitalization, on average, during our sample period) are not dramatic, they are remarkably robust. Continuation of the current exponential growth of ETF markets presents the potential for large increases return correlation levels, a negative externality for markets. Indeed, increased return correlation not only affects ETF investors but also investors who invest via traditional vehicles such as mutual funds and pension funds.

2 Data

2.1 ETF data

Although the first ETF began trading in 1980, Figure 1 shows that holdings of exchange-traded funds were a negligible percentage of stocks' shares outstanding prior to mid-2006, so our data begins in July of 2006. We obtain data on 1,413 exchange-traded funds from the CRSP stock database identified by their share code of 73. There are only 297 ETFs in CRSP in July of 2006, and this number steadily grows to 1,223 by the last month of our sample, June 2012. As ETFs are securities according to the CRSP stock database and funds according to the CRSP Survivor-Bias-Free Mutual Fund database, we can obtain both the fund's price information and its holdings information, which we match by cusip. We confirm that the funds are ETFs by retaining only funds with etf_flag of "F" in the CRSP mutual fund database. We further retain only equity ETFs, with Lipper asset code EQ in the CRSP mutual fund data base. In addition, we exclude foreign and global ETFs as described by their Lipper Class Name. ETF shares outstanding data are from Morningstar, which is more precise on a daily basis than the *shrout* variable from CRSP. When shares outstanding is missing in Morningstar, we use CRSP *shrout*.

We also obtain information on the stocks held by ETFs. We use the CRSP mutual fund holdings database because few ETFs are linked to the Thompson holdings database by the MFLinks linking database. Since portfolios are disclosed quarterly, on any given day the estimate of portfolio holdings is as of the latest quarterly disclosure multiplied by the number of shares outstanding today and divided by the number of shares outstanding at the time of disclosure.[8] Some funds, like Vanguard, use the same overall portfolio (crsp_portno) to disclose holdings by their mutual funds and ETFs together (various crsp_fundnos) Thus, for disclosure purposes, they treat the ETF as a separate share class of their traditional mutual fund. To capture only the ETF holdings, we use the assets under management in the ETFs and multiply by the percentages of the holdings in the overall portfolio. The median ETF in our sample turns over its portfolio only 0.25 times per year, which reflects that ETFs rarely change their portfolios. As such, holdings observed at the beginning of the quarter should measure the ETF portfolio composition during the quarter quite well. Our final sample consists of 699 US equity ETFs with holdings data. Consistent with the growth of the ETF sector during our sample period, the number of ETFs in our final sample grows from 152 in 2006 to 512 in 2012 (see Figure 1). Figure 1 also shows that ETF holdings for an average stock in our final sample grew fast as well.

2.2 Stock data

The study uses all stocks in CRSP with share codes 10 and 11 that have a market capitalization of 100 million dollars or more and a share price of $5 or more. Quarterly book values are from the Compustat database. Our final sample consists of 4,703 stocks. Among them, 4,318 stocks are held by at least one ETF during our sample period.

2.3 Summary statistics

Summary statistics on monthly data for ETFs in our sample appear in Table 1, Panel A. The average ETF in our sample holds about 0.1% of the total market capitalization of its underlying

[8] Using unadjusted holdings as of the latest quarterly disclosure does not change the nature or significance of our results.

portfolio. The median fund holds 0.011% of its underlying portfolio. Thus, 699 such ETFs add up to a nontrivial proportion of the market capitalization of the stocks they own. Moreover, the turnover of the funds, which averages 4.6% per day, is large compared to the turnover of the underlying stocks, which averages 1.1% per day (see Table 1, Panel B). If a stock is owned by many similar ETFs with turnover ratios of 4.6% per day, much of it due to arbitrage activity between the ETF and the underlying basket, one could see how stock prices could be affected.

The median total net assets (TNA) of the ETFs in our sample is 87 million but the average is larger, at 966 million. This is due to a few large ETFs such as State Street's SPY ETF, which tracks the S&P 500 index. The stock-level analysis includes a S&P 500 membership indicator variable in addition to stock fixed effects to ensure that the results are not due to S&P 500 membership and thus inclusion in some of these large ETFs. Consistent with ETFs being inexpensive to manage, expense ratios are very low, averaging half a percent per year. *N holdings* is the number of holdings of common stock in the fund's reported portfolio that can be matched to our stock sample. These funds hold an average of 240 stocks (median is 78) that pass the stock screens described above. We control for the log of the number of holdings in our later analysis since it can affect portfolio diversification.

Table 1, Panel B presents summary statistics for stocks in our sample. The mean and median ETF holdings of these stocks are more than 2%, comparable to the holdings by index funds. While the average ETF holding is small relative to that of the mutual funds (21.8%) and other institutions (44.5%), it has been growing exponentially in the recent past as evident in Figure 1. As a result, it is common for a stock to be held by multiple ETFs. In fact, the average stock in our sample is held by 27.6 ETFs and more than 25% of our sample stocks are each held by more than 41 ETFs. It is therefore crucial to include a broad cross-section of ETFs when measuring a stock's exposure to ETF activities.

3 ETF activities and return comovement

We first examine whether ETF activities are related to return comovement among component stocks at the fund level, and next we investigate whether ETF activities affect correlations with the market portfolio at the stock level.

3.1 Fund-level tests

In this subsection, we test whether an ETF's greater ownership of its underlying portfolio, creation and redemption activity, and turnover are related to the return correlations of its underlying stocks. We describe the measures of fund-level average return correlation and ETF activity below.

3.1.1 Empirical measures

We define the fund-level variance ratio ($Fratio$) as follows:

$$Fratio = \frac{\text{Variance of the average daily return of the stocks in the portfolio}}{\text{Average of the variances of the returns of stocks in the portfolio}}. \tag{1}$$

This ratio is computed each month for each ETF. According to equation (10) of Pollet and Wilson (2008), $Fratio$ is a proxy for average correlation among stocks in the portfolio. For intuition, consider an equal-weighted portfolio containing N stocks, the portfolio return variance during any period t, $\sigma^2_{p,t}$, is:

$$\begin{aligned}\sigma^2_{p,t} &= \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{1}{N^2}\rho_{jk,t}\sigma_{j,t}\sigma_{k,t} \\ &= \bar{\sigma}^2_t\bar{\rho}_t + \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{1}{N^2}\rho_{jk,t}\xi_{jk,t},\end{aligned} \tag{2}$$

where:

$$\bar{\sigma}_t^2 = \frac{1}{N}\sum_{j=1}^{N}\sigma_{j,t}^2, \tag{3}$$

$$\bar{\rho}_t = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\rho_{jk,t}, \tag{4}$$

$$\xi_{jk,t} = \sigma_{j,t}\sigma_{k,t} - \bar{\sigma}_t^2. \tag{5}$$

Pollet and Wilson (2008) show that the product between the average variance and the average correlation (the first term in the RHS of equation (2)) explains more than 97% of the variation in the portfolio return variance (the LHS of equation (2)). It then implies that the ratio between portfolio return variance and average stock return variance as in $Fratio$ should be the main driver of the average stock correlation. In fact, $Fratio$ is identical to the average stock correlation in the special case where all stocks have the same variance (so the second term in the RHS of equation (2) disappears). We winsorize $Fratio$ at the 1% level to remove the effect of outliers. Table 1, Panel A shows that *Fratio* has a mean of 0.46 and median of 0.44.

We use three measures of ETF activity at the portfolio level. The first measure is the proportion of the underlying portfolio that is held by the ETF, *Holdings%*. This is equal to the market capitalization of the ETF divided by the market capitalization of the underlying portfolio.

The second measure we use is the standard deviation of the daily number of shares outstanding of the ETF, divided by the mean shares outstanding during the month, *SD shares*. This is meant to capture the intensity of the creation and redemption activity of the ETF. Creation and redemption could drive underlying stock correlations if authorized participants (APs) need to buy and sell large parts of the portfolio together when they create or redeem shares, but since creation and redemption occurs only once a day and is thus unlikely to be used in arbitrage, there is less urgency for the entire portfolio to trade together. Panel A of Table 1 shows that the daily standard deviation of shares outstanding averages 2.8% of ETF shares outstanding per day. The median is smaller at

1%.

A third measure of ETF activity is *ETF turnover*. This is the average over the month of the ratio of the daily number of shares traded to the number of shares outstanding that day. This will be positively related to the amount of arbitrage activity in the ETF, although there are clearly other reasons to trade the ETF besides arbitrage of its price relative to its components' prices. Table 1, Panel A shows that ETF turnover averages 4.6% per day.

We use fund and time fixed effects, which subsume many possible control variables such as industry classification, market return and volatility. However, some controls vary by both fund and time. We include fund size as measured by total net assets (TNA). We also include the number of holdings as a control variable since it can affect portfolio diversification and thus the $Fratio$.

ETFs often hold stocks in common, so regression errors may be correlated in the cross-section. As a result, standard errors that are double-clustered by month and fund, and thus robust to heteroskedasticity and autocorrelation, are not sufficient for our purposes. We follow Driscoll and Kraay (1998) to compute a nonparametric covariance matrix estimator that produces standard errors that are robust to general forms of spatial and temporal dependence.[9] The Driscoll-Kraay standard errors are consistently larger than the (untabulated) double-clustered standard errors in our panel regressions.

3.1.2 Empirical results

In Table 2, we regress the measure of in-portfolio correlation, *Fratio*, on *Holdings%*, *SD shares*, *ETF turnover* and control variables. Panel A presents univariate regressions showing that all three measures of ETF activity are positively related to *Fratio*. Panel B presents the regressions with time and fund fixed effects and control variables. Columns 1-4 have only time fixed effects and columns 5-8 have both time and fund fixed effects. When only time fixed effects are included, *Holdings%* and *SD shares* are still significant (columns 1-2), but they are no longer significant when fund fixed

[9]The Driscoll-Kraay standard error is estimated in Stata using the xtscc program prepared by Driscoll and Kraay (1998).

effects are added. In contrast, *ETF turnover* is significant in all regression specifications.

In column 8, all three explanatory variables appear together along with time and fund fixed effects. This column shows that the strongest predictor of how much the stocks in the portfolio co-move is the daily turnover of the ETF. A one-standard-deviation increase (0.105) in the daily turnover of an average ETF in our sample is associated with a 0.113*0.105=0.012 increase in the *Fratio* of the stocks in its portfolio. This amounts to 6.6% of its standard deviation.

We note that *Holdings%* and *SD shares* are no longer significant when time and fund fixed effects are included. The result helps to make two points. First, ETF arbitrage as proxied by *ETF turnover* most likely drives return correlations. Creation and redemption activity is less important, which is not surprising. Creation and redemption is more onerous than trading and typically occurs after market close, when an arbitrage opportunity may no longer be available. Second, *Holdings%* and *SD shares* may also proxy for investors time varying style preference. For example, an increase in investors' interest in value stocks may result in creations of new shares of ETFs specializing in value stocks, thus lead to increases in both *Holdings%* and *SD shares*. The fact that *ETF turnover* remains highly significant after controlling for *Holdings%* and *SD shares* suggests that the increased return correlation most likely come from ETF arbitrage rather than time-varying style investment.

The impact of ETF activity on underlying stock return correlations is not equal across ETFs. Larger ETFs, by holding bigger fractions of the total market capitalization of their underlying stocks, could drive the stock correlations more. To test this conjecture, we sort our cross-section of ETFs into three subsets based on total net asset value and repeat our regressions in Table 2, Panel B in each subset. Indeed, in Table 3, Panel A, we find ETF turnover to have the largest coefficient (0.137) among the largest ETFs. Among these largest ETFs, even *SD shares*, which proxies for creation and redemption activity, is significant.

It is important to note that ETF turnover on its own should not generate higher stock correlations. ETF turnover affects stock correlations only insofar as it proxies for equivalent turnover in the underlying stocks via arbitrage trades. In such arbitrage trades, ETF turnover and the underlying stock turnover should occur simultaneously, which motivates our second subsample cut. Each month, we regress each ETF's daily turnover on the average daily turnover of its underlying stocks and compute the R^2 of the regression. A high R^2 indicates more simultaneous trading in both the ETF and its underlying stocks, such that *ETF turnover* more likely reflects arbitrage trading and affects stock correlations. Table 3, Panel B confirms that *ETF turnover* has a higher coefficient (0.115) among the tercile of ETFs with the highest R^2s.

Last, we would expect an arbitrage-driven increase in correlation to be smaller for ETFs with illiquid holdings. While these are equity ETFs, some hold a small proportion of other assets, such as corporate or municipal bonds, preferred stock, asset-backed securities or mortgage-backed securities in addition to stocks and cash-equivalents. Column 2 of Table 3, Panel C examines the subset of 121 ETFs that holds any of these relatively less liquid assets. Column 3 of Table 3, Panel C examines the sample where intraday arbitrage should be relatively easier - when the proportion of common stock is greater than the sample median of 99.67%. In this subset, we find that the coefficient on *ETF turnover* is larger than in the full sample presented in column 1. The last row of this table presents the average absolute value of the gap between NAV and closing price during the sample period. NAV is from the CRSP mutual fund database daily file and closing prices are from the CRSP stock database. This row shows that such a gap is greatest in the low-liquidity sample and lower in the third column when the ETF is mostly stock, compared to the full sample value of 0.25%. This pattern suggests that a large gap may actually reflect the difficulty of arbitrage rather than an opportunity for arbitrage.

3.2 Stock-level tests

In this sub-section, we focus our attention on stock-level analysis. If arbitrage between an ETF and its component stocks results in trading of that basket of stocks, we would expect a stock with greater exposures to ETF activities to experience more correlated trading and to comove more with the market. We test this prediction using stock-level data.

3.2.1 Empirical measures

For these stock-level tests, we need more generalized measures of comovement because stocks are held by many different ETFs. Thus, our measures relate to comovement with the market as a whole. As in Bekaert, Hodrick, and Zhang (2009) equations (3) and (4), these measures will be positively related to average measures of pairwise comovement between stocks. We use two measures of how much a stock comoves with the market. The first measure is the CAPM beta, β_M, as commonly studied in the comovement literature. This is the beta obtained from a regression of excess stock returns on the excess market returns provided by Kenneth French's Web site.

To account for potential comovement of the stock with other portfolio-based factors in addition to the market, we consider a second measure defined as the total variance of the stock divided by the idiosyncratic variance. This number will be high if the stock comoves with multiple factors and thus does not have much idiosyncratic variance. We call this ratio *Sratio*:

$$Sratio = \frac{\text{Total variance of the stock's returns}}{\text{Idiosyncratic variance of the stock's returns}}. \tag{6}$$

The variance of the stock's returns is the variance of of daily returns taken over the month, and the idiosyncratic variance is the variance of daily Fama and French (1993)+ Carhart (1997) 4-factor model adjusted returns. To adjust the returns, we calculate the coefficients on the Fama and French and Carhart factors with a 5-year rolling window of daily data prior to each date. The risk-free rate and market returns are from Kenneth French's website. Months with less than 15 observations are removed. *Sratio* is similar in spirit to the *Fratio* we used at the fund level. As with the *Fratio*, we winsorize *Sratio* at the 1% level to remove the effect of outliers. Summary statistics of these variables appear in Table 1, Panel B.[10]

We have three measures of ETF activity at the stock level. The first measure, *ETF%*, is the

[10] We also consider additional measures that are designed to capture the correlations between a stock and aggregate factors. Since results are almost identical, we leave them untabulated.

proportion (in percentage) of the stock's market capitalization that is held by all ETFs in our sample. This is computed using the holdings data in the CRSP mutual fund database. The second measure, *Wtd SD*, is the weighted average (by the proportion of the stock they hold) of the standard deviation of shares outstanding of the ETFs holdings the stock.

$$Wtd\ SD_{i,t} = \frac{\sum_{j=1}^{N} w_{i,j,t} SD\ shares_{j,t}}{\sum_{j=1}^{N} w_{i,j,t}}, \tag{7}$$

where j indexes the ETF, i indexes the stock, $w_{i,j,t}$ is the weight held by ETF j in stock i at time t, and N is the number of ETFs holdings the stock:

The third measure of ETF activity, *Wtd turnover*, is the weighted average of the turnover of the ETFs holding the stock.

$$Wtd\ turnover_{i,t} = \frac{\sum_{j=1}^{N} w_{i,j,t} ETF\ turnover_{j,t}}{\sum_{j=1}^{N} w_{i,j,t}}. \tag{8}$$

Here again, j indexes the ETF, i indexes the stock, $w_{i,j,t}$ is the weight held by ETF j in stock i at time t, and N is the number of ETFs holdings the stock.

Although we will use stock and time fixed effects, some time-varying firm-level control variables are also included in the regressions. Summary statistics for these variables appear in Panel B of Table 1. We include the average daily turnover of the stock, which is its volume from CRSP (vol) divided by its shares outstanding *(shrout*1,000)*. We also include the log of the stock's market capitalization from CRSP and its book/market ratio(B/M) ratio, where the denominator is the market capitalization and the numerator is the latest reported book value from Compustat. S&P 500 is an indicator variable for whether the stock belongs to the S&P 500 index in that month. *Stock turnover* is the stock's average daily turnover during the month. *Index%*, *MF%* and *Ins%* are total index fund holdings, mutual fund holdings and total institutional holdings in percentages, respectively. Mutual fund and index holdings are computed using the CRSP mutual fund holdings database, and institutional holdings are computed using the Thomson database of quarterly holdings. Since ETFs and index funds are mutual funds, their holdings are subtracted from total mutual fund holdings. Institutional holdings are computed using all categories of institutions in the Thomson institutional database and subtracting total CRSP mutual fund holdings, index fund and ETF holdings. The most recent holdings prior to the end of the current quarter are used. In contrast, we use ETF holdings reported as of the end of the latest month to mitigate endogeneity concerns.

Flow-induced trading by mutual funds can be another reason for higher return comovement as the funds often scale up or down their entire portfolios to satisfy the flow. Lou (2012) proposes a measure to capture such flow-induced price pressure (FIPP) based on mutual fund holdings. A stock associated with higher FIPP (in absolute terms) arguably will be associated with higher return comovement as it is often traded together with a portfolio of other stocks. We find that controlling for Lou's *Abs(FIPP)* variable in our analysis does not drive out the significant coefficients on our measures of ETF activities, suggesting that the ETF-arbitrage-induced return comovement is distinct from that arising from flow-induced trading.[11]

3.2.2 Comovement with the market

Table 4 presents regressions at the stock and month level of measures of how the stock covaries with the market on measures of ETF holdings and activity. Panel A presents univariate results and results where just the three explanatory variables are in the regression together. Driscoll-Kraay standard errors appear in these tables since these are more conservative than double-clustered standard errors. Panel A shows that all three explanatory variables are related to both measures of comovement. When they are together, however, not all variables remain statistically significant. The proportion of the stock that is held by ETFs predicts comovement with the market in the subsequent month. As in the fund-level tests, the strongest correlate of comovement of a stock

[11]We thank Dong Lou for supplying us with his FIPP data.

with the market portfolio seems to be weighted average turnover of the ETFs holding the stock, and the measure of creation and redemption is the weakest. The coefficient on *Wtd SD* sometimes changes sign when it is included with the other variables.

Panels B and C present multivariate results when both time and stock fixed effects are included. Panel B presents β_M as the dependent variable; Panel C presents *Sratio*. All coefficients on both explanatory variables of interest are lower. Importantly, the effect of ETF holdings is more than three times the effect of mutual fund holdings or of institutional holdings. In column 1, Panel B, a 1% increase in ETF holdings of a stock is associated with a 0.027 increase in its CAPM beta. This is not a large increase compared to β_M's mean value of 1.183 and standard deviation of 0.755, but if ETFs become comparable in size to mutual funds, which have around 20% ownership of many stocks, the associated increase in β_M could be much larger.

Wtd SD is significant on its own with fixed effects and controls (see column 2 of Panels B and C), but when it is with the other two explanatory variables of interest, its sign changes (column 4 of Panel B), although its effect remains positive and significant in Panel C. This mirrors the weaker performance of this variable in the fund-level tests in the prior section. Therefore, we do not consider it as reliable a driver of correlation as *ETF%* or *Wtd turnover*.

Wtd turnover is significantly positively related to all three measures of comovement regardless of the other variables in the model. Table 4, Panel B, column 4 shows that a one-standard-deviation increase in *Wtd turnover* is associated with a 0.720*0.112=.08064 increase in β_M.

The effects are similar with *Sratio*. In each case, the effect of a percentage increase in ETF holdings is much larger than the effect of a percentage increase in mutual fund holdings or institutional holdings. Column 4 of Panel C shows that a one-standard-deviation increase in *Wtd turnover* is associated with an 0.824*0.112=0.092 increase in *Sratio*. Indeed, when we later run cross-sectional regressions of a measure of stock comovement with the market (*Sratio* or β_M) on *Wtd turnover* and other controls every month and plot the coefficients over time in Figure 2, we find these coefficients to be almost always positive, confirming the strong and persistent economic link between ETF turnover and return comovement.

3.2.3 Sub-sample cuts

Table 5 shows Table 4's results broken down by terciles of size and turnover. Only the coefficients on the variables of interest are shown but each regression also contains time and stock fixed effects and all of the control variables in Table 4. These tables show that the results tend to be stronger for the smaller stocks (Size 1). Recall that stocks with prices below \$5 or market capitalization below \$100 million are excluded, so these are not micro-cap stocks. Panel B shows that the effect is also strongest for stocks with lower turnover. These tests help shine a light on why *Wtd SD*, our proxy for creation and redemption activity, is less related to underlying asset correlations than turnover. This variable is robustly positively related to correlations in the smallest terciles of size and turnover but the effect is weaker in the largest tercile. For smaller and lower turnover stocks, it must be more difficult to locate or sell components during creation and redemption.

3.2.4 Time-series evidence

We present time-series evidence that further links arbitrage activities to return comovement. If excessive return comovement comes from non-fundamental shocks made systematic by ETF arbitrage, we would expect our results to be stronger during periods of turbulent markets when such shocks are plentiful. We use the Chicago Board Options Exchange Market Volatility Index (VIX) as a measure of market turbulence. Anecdotal evidence suggests a close link between the VIX and arbitrage opportunities measured by the profitability of the high frequency traders who are the natural ETF arbitrageurs.[12]

Each month, we run cross-sectional regressions of β_M or $Sratio$ on weighted average turnover of the ETFs (*Wtd turnover*) and other controls. The regression coefficient on *Wtd turnover* identify

[12] According to Rosenblatt's statistics, "with the CBOE Volatility Index having fallen from near 80 at the height of the credit crisis in October 2008 to nearly 12 at the outset 2013, conditions have gotten harsher for high-frequency trading firms, he said. Such firms' profits from equities trading have fallen 75%, from 2009."

the marginal impact of ETF trading on return comovement in that month. While average return comovement tends to be high during volatile periods, it is important to note that our regression coefficients are measuring the effect of ETF trading on return comovement from the cross-section. Figure 2 plots the lagged VIX index against these regression coefficients on β_M (top) or on $Sratio$ (bottom). We observe positive and significant correlation coefficients, almost 30% on average, suggesting that ETF trading has a bigger impact on return comovement precisely during periods with ample mispricings.

4 Is the return comovement excessive?

So far, we have provided evidence suggesting that ETF activity is positively related to return comovement. A natural question follows: does the increased return comovement reflect faster incorporation of common information in the market that ETF trading helps to facilitate; or does it reflect "excessive" price movement due to non-fundamental shocks that ETF trading helps to propagate? We address this important question in this section.

4.1 Fund-level test: correlation of largest two stocks and shocks to the other stocks

We first examine the effect of an exogenous, non-fundamental shock to some of the stocks in the portfolio has on the remaining stocks. Flow-induced price pressure, coined by Lou (2012), is an example of such a non-fundamental shock. Our hypothesis is that if some stocks in an ETF are pushed in one direction by flow-induced price pressure (FIPP), resulting in a wedge between the ETF price and its underlying value, subsequent ETF arbitrage will increase correlations among all stocks, including those that are not directly affected by the initial flow-induced price pressure. For a clean test, we select two stocks - the two stocks with the largest portfolio weights in the ETF - and calculate the weighted average FIPP of the remaining stocks. In other words, the independent variable is not the FIPP on the two largest stocks themselves, but the weighted average FIPP on the stocks connected to them through the ownership of the ETF. We choose the largest stocks because they must be large enough in weight to be traded by an arbitrageur wishing to profit from discrepancies in the price of the ETF and the underlying portfolio, and because they should be large enough to affect the price of the ETF.

We regress the monthly correlation of daily returns of stocks 1 and 2 on the absolute value of the weighted FIPP of the remaining stocks. Stock-level control variables are the log of the market capitalization, the weight of the stock in the ETF, and the absolute value of the stock's FIPP. Fund-level control variables include *Expense ratio*, number of holdings: *N holdings*, and *Log(TNA)*, the log of the fund's TNA. We also have fund and time fixed effects. Summary statistics of all variables appear in Table 1, Panel A. Results of the regressions of $Corr(R_1, R_2)$ on Abs(Wtd $FIPP_{\{1,2\}^{\perp}}$) appear in Table 6. In the last column which includes all of the control variables, a one-standard-deviation increase in *Abs(Wtd* $FIPP_{\{1,2\}^{\perp}}$*)* is associated with an 0.01 increase in $Corr(R_1, R_2)$. This is 3.33% of a standard deviation in $Corr(R_1, R_2)$. Recall that this is the effect of just one ETF on two of its component stocks.

4.2 Stock-level test: lagged betas

We next examine a stock's lagged market betas. If an individual stock return on day t contains a component that reflects "excessive" comovement, such a component is likely to revert in the next two days. As a result of this reversal, stock returns on day $t+1$ and $t+2$ will likely load negatively on the market return on day t. In other words, the stock will have negative lagged market betas. In contrast, if higher return comovement comes from faster incorporation of common information, we should not observe return reversals and hence lagged betas should never be negative.

For each stock during our sample period and in each month, we compute betas of stock returns on contemporaneous and lagged daily market returns, as follows:

$$R_{i,t} = \sum_{l=0}^{4} \beta_{R_M,i,l} R_{M,t-l} + \epsilon_{i,t} \tag{9}$$

where $R_{M,t}$ is the return on the market portfolio and $R_{i,t}$ is the daily return on stock i, and l is the lag. Note that contemporaneous market returns are included by $l = 0$. We winsorize all betas at the 1% level to mitigate the effect of outliers. Table 7 examines means of these betas by tercile of measures of ETF activity, and compares the third (highest) tercile of ETF activity to the first (lowest) tercile of ETF activity. Table 7 shows that the lowest tercile of ETF activity tends to have postive betas on lagged market returns, while stocks in the highest tercile of ETF activity tend to have negative betas on lagged market returns. This suggests that stocks in the lowest tercile of ETF activity tend to be delayed, while stocks in the highest tercile tend to overshoot and show reversals, suggesting "excess" comovement. As in the prior tables of the paper, the results for *Wtd turn* are the strongest. The results are strongest for lags 1 and 2. They are not significant for lags 3 and 4 of market returns and those results remain untabulated.

This effect is confirmed in Table 8, in a regression controlling for other potential determinants of the betas on lagged market returns. This regression shows that measures of ETF activity are negatively related to betas on lagged market returns. As in Table 7, the results are strongest for *Wtd turn*.

5 Conclusion

We provide empirical evidence that the arbitrage activity between an Exchange-traded fund (ETF) and its underlying portfolio propagates non-fundamental shocks from the ETFs to a broad cross-section of stocks they hold. We first perform an ETF-level analysis and find that an ETF's turnover is a strong determinant of the comovement of the stocks in its portfolio. This holds controlling for fund and time fixed effects and a host of fund-level control variables. In addition, the more an ETF owns of the market capitalization of its underlying portfolio, the more the stocks in that portfolio tend to move together in the subsequent month. This relation, however, becomes insignificant after controlling for fund and time fixed effects and a host of control variables. Finally, the standard deviation of the ETF's daily shares outstanding, capturing creation and redemption activity, is less strongly related to comovement. This is possibly because it is quite low relative to turnover, and due to the constraints of creation and redemption, likely not as closely related to the synchronized trading of the underlying portfolio.

Next, we investigate how ETF ownership of a stock in aggregate relates to its comovement with the market portfolio. We find that the more of a stocks' market cap is owned by ETFs, the more it comoves with the market in the subsequent month. As in the fund-level analysis, we also find that the weighted average turnover of the ETFs that own the stock is related to how much the stock comoves with the market. We find little evidence that weighted average creation and redemption activity of the ETFs that hold the stock is related to the stock's comovement with the market.

We find evidence that this increased comovement is more likely excessive than reflecting an increased efficiency in incorporating common shocks. First, at the fund level, we find that the correlation between the two largest stocks increases with the flow-induced price pressure of the remaining stocks, suggesting that the associated ETF mispricing and resulting arbitrage push up stock correlations. Next, we find that sensitivity to lags of market returns is more negative for stocks with high ETF activity, controlling for other possible drivers of this relationship. This suggests that ETF activity is related to simultaneous overshooting and reversals in prices, a sign of "excessive" comovement.

There is no doubt that the ETF structure provides great benefits. Among others, ETFs provide a cheaper and more efficient way for investors to diversify into a broad asset portfolio. At the same time, the results in our paper suggest that they may also lead to "excessive" comovement among these assets that could reduce this diversification benefit. With the exponential growth and popularity of ETFs, our findings are clearly relevant for developing the ETF market going forward.

References

Abner, David, 2010, *The ETF handbook* (John Wiley and Sons, Inc., Hoboken, New Jersey.).

Anton, Miguel, and Christopher Polk, 2013, Connected stocks, *Journal of Finance* forthcoming.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

——— , and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283–317.

Bartram, Sohnke M., John Griffin, Tae-Hoon Lim, and David T. Ng, 2013, How important are foreign ownership linkages for international stock returns?, *Working paper*.

Bekaert, Geert, Robert Hodrick, and Xiaoyan Zhang, 2009, International stock return comovements, *Journal of finance* 64, 2591–2626.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2012, Etfs, arbitrage, and contagion, *Working paper*.

Boehmer, Beatrice, and Ekkehart Boehmer, 2003, Trading your neighbor's ETFs: Competition or fragmentation?, *Journal of Banking and Finance* 27.

Bradley, Harold, and Robert E. Litan, 2011, ETFs and the present danger to capital formation, *Testimony before the Subcommittee on Securities, Insurance, and Investments of the Senate Banking Committee*.

Driscoll, John C., and Aart C. Kraay, 1998, Consistent covariance matrix estimation with spatially dependent data, *Review of Economics and Statistics* 80, 549–560.

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2012, The real effects of financial markets: The impact of prices on takeovers, *The Journal of Finance* 67, 933–971.

Engle, Robert, and Debojyoti Sarkar, 2002, Pricing exchange traded funds, *Working paper*.

Forbes, Kristin J., and Roberto Rigobon, 2002, No contagion, only interdependence: Measuring stock market comovements, *The Journal of Finance* 57, 2223–2261.

Gagnon, Louis, and G. Andrew Karolyi, 2010, Multi-market trading and arbitrage, *Journal of Financial Economics* 97, 53–80.

Goetzmann, William, and Massimo Massa, 2003, Index funds and stock market growth, *Journal of Business* 76, 1–28.

Greenwood, Robin, 2008, Excess comovement of stock returns: Evidence from cross-sectional variation in nikkei 225 weights, *Review of Financial Studies* 21, 1153–1186.

——— , and N. Sosner, 2007, Trading patterns and excess comovement of stock returns, *Financial Analysts Journal* 63, 69–81.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Hamm, Sophia, 2011, The effect of ETFs on stock liquidity, *Working paper.*

Harris, Larry, and E. Gurel, 1986, Price and volume effects associated with changes in the S&P 500: new evidence for the existence of price pressures., *Journal of Finance* 41, 815–829.

Hong, Harrison, Jeffrey D. Kubik, and Tal Fishman, 2012, Do arbitrageurs amplify economic shocks?, *Journal of Financial Economics* 103.

Jiang, Wenxi, and Hongjun Yan, 2012, Financial innovation, investor behavior, and arbitrage: Implications from levered etfs, *Working paper.*

Kaul, Aditya, Vikas Mehrotra, and Randall Morck, 2002, Demand curves for stocks do slope down: new evidence from an index weights adjustment, *Journal of Finance* 55, 893–912.

Khan, Mozaffar, Leonid Kogan, and George Serafeim, 2012, Mutual fund trading pressure: Firm-level stock price impact and timing of seos, *The Journal of Finance* 67, 1371–1395.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457–3489.

——— , and Christopher Polk, 2013, Comomentum: Inferring arbitrage activity from return correlations, *Working paper.*

Lynch, Anthony, and Richard Mendenhall, 1997, New evidence on stock price effects associated with changes in the s&p 500 index, *Journal of business* 70, 351–383.

Marshall, Ben, Nhut Nguyen, and Nuttawat Visaltanachoti, 2012, The microstructure of arbitrage: ETF evidence, *Working paper.*

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.

Petajisto, Antti, 2011, Inefficiencies in the pricing of exchange-traded funds, *Working paper.*

Pollet, Joshua M., and Mungo Wilson, 2008, How does size affect mutual fund behavior?, *Journal of Finance* 63, 2941–2969.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *Journal of Finance* 41.

——— , and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

Staer, Arsenio, 2012, Equivalent volume and comovement, *Working paper.*

Sullivan, Rodney N., and James X. Xiong, 2012, How index trading increases market vulnerability, *Financial Analyst Journal* 68, 70–84.

Wurgler, Jeffrey, 2010, *Challenges to Business in the Twenty-First Century: The Way Forward* . chap. On the Economic Consequences of Index-Linked Investing (American Academy of Arts and Sciences).

——— , and E. Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks?, *Journal of Business* 75.

Appendix

Variable definitions

Variable	**Definition**
	ETF-level measures
ETF turnover	Average daily turnover of ETF shares. Turnover is volume (CRSP vol) divided by shares outstanding from Morningstar. CRSP shares oustanding is used if Morningstar shares outstanding is missing.
Expense ratio	Expense ratio of the fund, from CRSP mutual fund database, in percent.
Fratio	The ratio of the variance of the portfolio to the average of the variances of the stocks in the portfolio. The ratio is winsorized at the 1% level.
Holdings %	The proportion of the portfolio's total market capitalization that is owned by the ETF on the last day of the prior month. Holdings are computed using the CRSP Survivor-Bias-Free Mutual fund holidngs database.
Log(N holdings)	The log of the number of holdings of common stock reported to CRSP Mutual fund database and matched to CRSP stock database, that have prices above \$5 and market capitalizations above \$100,000,000.
SD shares	The standard deviation of ETF shares outstanding, from Morningstar. CRSP shares outstanding is used if Morningstar shares outstanding is missing.
TNA	Total net assets of the fund as of the latest report, from the CRSP Mutual Fund Database, in millions of dollars.
$Corr(R_1, R_2)$	The monthly correlation between the daily returns of the two stocks with the largest portfolio weights in the fund.
Abs(Wtd $FIPP_{\{1,2\}^{\perp}}$)	The absolute value of the weighted average FIPP of the third through last stock in the fund in terms of portfolio weights.
Abs($FIPP_1$)	The absolute value of the FIPP of the fund holding with the largest portfolio weight.
Abs($FIPP_2$)	The absolute value of the FIPP of the fund holding with the second-largest portfolio weight.
Log(Marketcap_1)	The log of the market capitalization of the fund's holding with the largest portfolio weight.
Log(Marketcap_2)	The log of the market capitalization of the fund's holding with the second-largest portfolio weight.
Percent TNA_1	The fund's largest portfolio weight.
Percent TNA_2	The fund's second-largest portfolio weight.
	Stock-level measures
β_M	Coefficient of the stock's daily excess returns on daily market excess returns in that month.
B/M	Book to market ratio. Book values are from Compustat and market value is log(abs(prc)*shrout*1000) from CRSP. Negative book values are set to missing.

ETF %	Proportion of the stock that is held by ETFs, using CRSP mutual fund holdings data and market capitalization data, on the last day of the prior month.
Ins. %	Percentage of the stock that is held by any institution in the Thomson database minus the proportion that is held by CRSP mutual funds.
Log(Mkt cap)	The log of the firm's market capitalization, $log(abs(prc) * shrout * 1000)$, from CRSP.
MF %	Percentage of the stock that his held by CRSP mutual funds minus the proportion that is held by ETFs.
N ETF holders	The number of ETFs holding the stock.
S&P 500	Indicator for whether the stock is a S&P 500 member in that month.
Sratio	Ratio of stock variance to stock idiosyncratic variance of daily returns over the month. The ratio is winsorized at the 1% level.
Stock turnover	Average daily turnover over the month from CRSP. Turnover is $vol/(shrout * 1000)$.
Wtd SD	Weighted average percentage standard deviation in the shares outstanding of the ETFs that hold the stock. The weights are proportional to the holdings of each ETF that holds the stock.
Wtd turnover	Weighted average turnover of the ETFs that hold the stock. The weights are proportional to the holdings of each ETF that holds the stock.

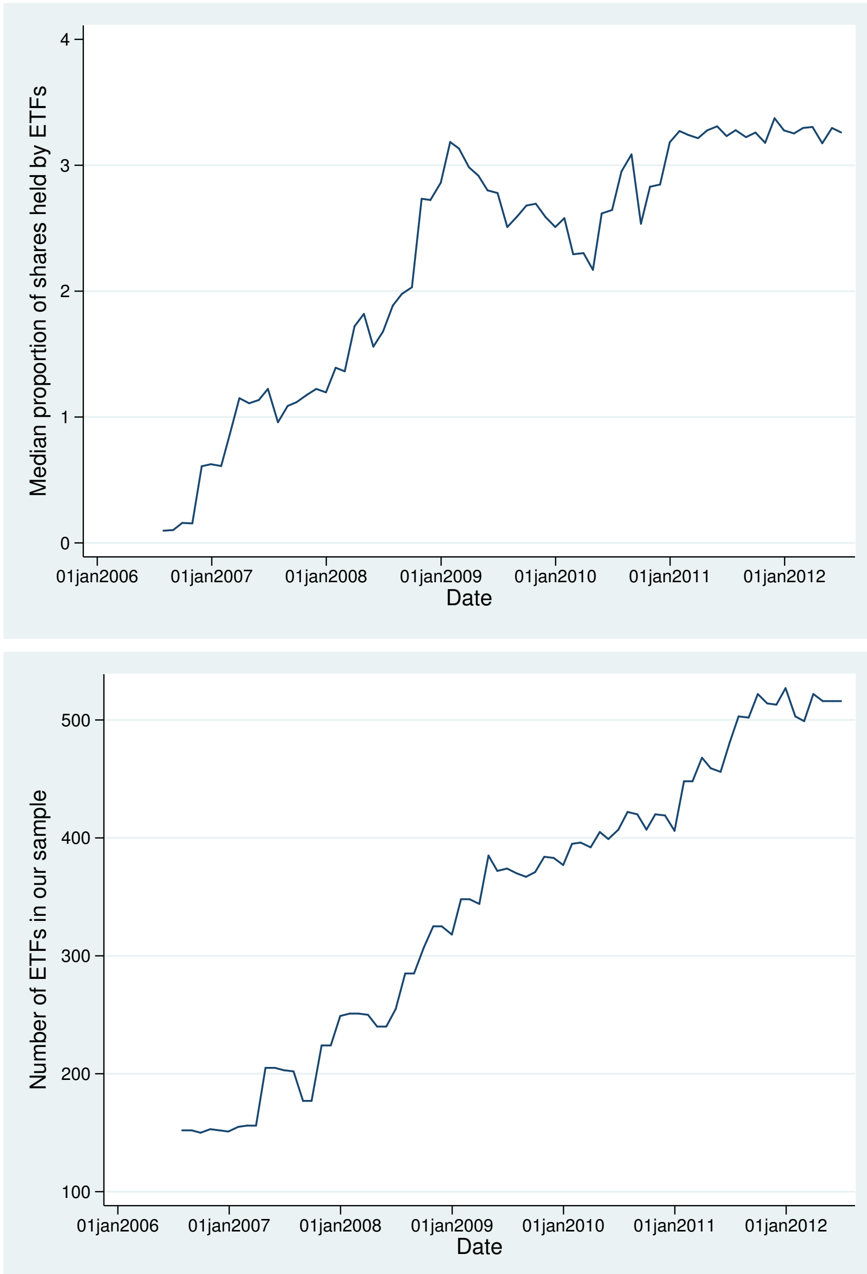

Figure 1. The top figure presents the median of the percentage of the stock that is held by exchange traded funds for CRSP stocks with share price of at least \$5 and market capitalization of at least \$100 million. The bottom figure presents the number of ETFs in our sample each month. The sample consists of purely domestic equity ETFs.

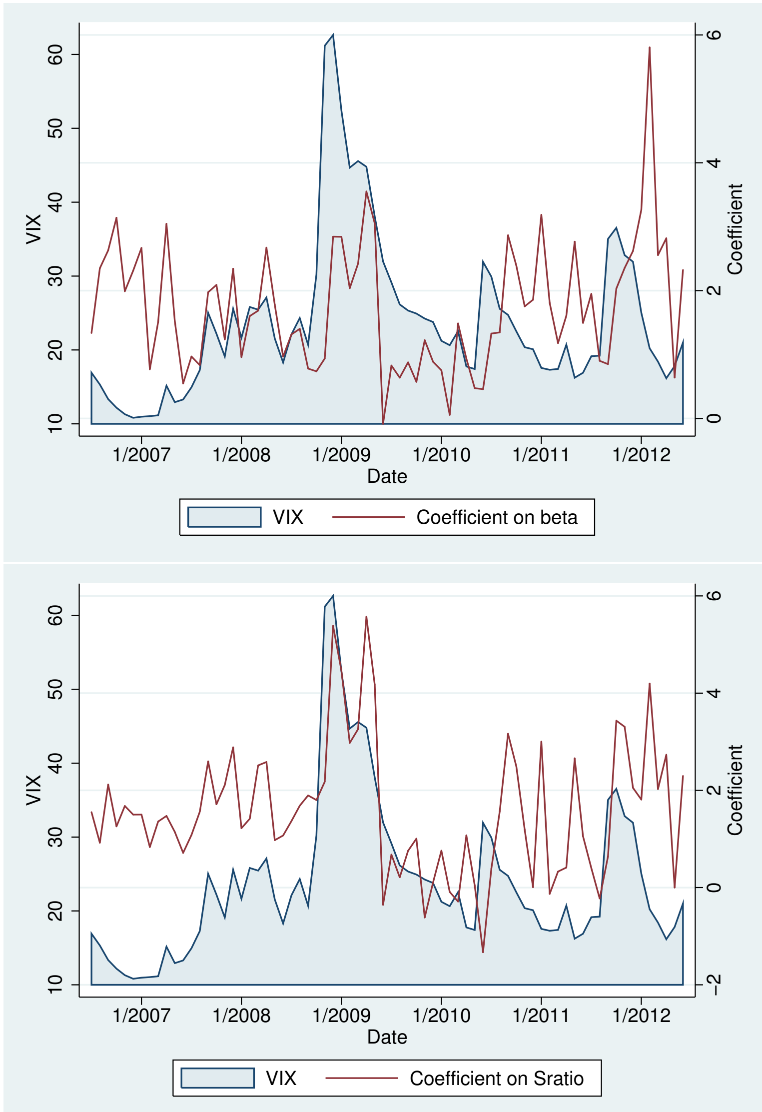

Figure 2. Coefficients of stock-level measures of correlation with the market, β_M (top), and *Sratio* (bottom) regressed on ETF weighted average turnover of those stocks, *Wtd turn*. β_M is the regression coefficient of the stock's daily excess returns on daily market excess returns in that month. *Sratio* is the ratio of stock variance to stock idiosyncratic variance of daily returns over the month. One-month lagged VIX also appears in each figure.

Table 1.

Summary statistics. Panel A presents summary statistics of monthly ETF-level data for 699 ETFs. $Fratio$ is the ratio of the variance of the portfolio to the average of the variances of the stocks in the portfolio. *Holdings %* is the proportion of the portfolio's total market capitalization that is owned by the ETF on the last day of the prior month. *SD shares* is the standard deviation of ETF shares outstanding. *ETF turnover* is the average daily turnover of ETF shares. *Expense ratio* is the annual expense ratio of the fund, in percent. *TNA* is total net assets of the fund as of the latest report, in millions of dollars. *N holdings* is the number of the ETF's holdings of common stock that are also in our CRSP stock sample. $Corr(R_1, R_2)$ is the monthly correlation between the daily returns of the two largest stocks in the fund. *Abs(Wtd* $FIPP_{\{1,2\}^{\perp}}$*)* is the absolute value of the weighted average FIPP of the third through last stock in the fund in terms of portfolio weights. *Abs(*$FIPP_1$*)* and Abs($FIPP_2$) are the absolute value of the FIPP of the fund holding with the largest and second-largest portfolio weights. *Log(*$Marketcap_1$*)* and *Log(*$Marketcap_2$*)* are the log of the market capitalization of the fund's holding with the largest and second-largest portfolio weights. *Percent* TNA_1 and *Percent* TNA_2 are the fund's largest and second-largest portfolio weights.

Panel B presents stock-level summary statistics for 4,703 stocks. *Sratio* is the ratio of stock variance to stock idiosyncratic variance of daily returns over the month. β_M is the coefficient of the stock's daily excess returns on daily market excess returns in that month. *ETF %* is the proportion of the stock that is held by ETFs, using CRSP data, on the last day of the prior month. *Wtd SD* is the weighted average percentage standard deviation in the shares outstanding of the ETFs that hold the stock. *Wtd turnover* is the weighted average turnover of the ETFs that hold the stock. *Stock turnover* is the average daily turnover of the stock over the month *Log(Mkt cap)* is the log of the firm's market capitalization. *B/M* is the book-to-market ratio. *S&P 500* is an indicator variable for whether the stock is currently in the S&P 500 index. Index % is the percentage of stock held by index funds. *MF %* and *Ins.%* are the percentage of the stock held by mutual funds and by other institutions. *N ETF holders* is the number of ETFs that hold the stock. *Abs(FIPP)* is the absolute value of the FIPP of the stock. Detailed variable definitions appear in the Appendix.

Panel A

Variable	Mean	SD	p1	p25	p50	p75	p99	N
Fratio	0.456	0.180	0.134	0.312	0.438	0.588	0.868	24,799
Holdings %	0.093	0.217	0.000	0.002	0.011	0.058	1.228	24,799
SD shares	0.028	0.048	0	0	0.010	0.033	0.283	24,799
ETF turnover	0.046	0.105	0	0.006	0.012	0.028	0.655	24,799
Expense ratio	0.005	0.003	0.001	0.003	0.005	0.007	0.010	21,112
TNA	966	4,622	2	20	87	364	13,137	23,950
N holdings	240	414	3	30	78	255	1942	24,799
$Corr(R_1, R_2)$	0.494	0.301	-0.347	0.302	0.535	0.732	0.954	24,760
Abs(Wtd $FIPP_{\{1,2\}^{\perp}}$)	0.051	0.096	0.000	0.010	0.022	0.043	0.600	24,815
Abs($FIPP_1$)	0.062	0.128	0.001	0.014	0.031	0.058	0.986	23,918
Abs($FIPP_2$)	0.067	0.144	0.001	0.015	0.032	0.060	1.032	23,756
Log(Market cap$_1$)	23.896	1.994	19.317	22.368	23.871	25.839	27.015	24,763
Log(Market cap$_2$)	23.709	1.876	19.307	22.285	23.711	25.496	26.734	24,726
Percent TNA$_1$	0.068	0.126	0.003	0.022	0.045	0.086	0.406	23,915
Percent TNA$_2$	0.049	0.085	0.002	0.018	0.036	0.065	0.243	23,877

Panel B

Variable	Mean	SD	p1	p25	p50	p75	p99	N
β_M	1.183	0.755	-0.620	0.701	1.125	1.610	3.527	195,490
Sratio	2.592	1.789	1.036	1.457	1.987	3.018	11.129	195,490
ETF %	2.414	1.954	0	0.818	2.147	3.549	8.591	195,490
Wtd SD	0.031	0.024	0	0.015	0.025	0.041	0.117	195,490
Wtd turnover	0.127	0.112	0	0.028	0.106	0.186	0.497	195,490
Stock turnover	0.011	0.016	0.000	0.004	0.007	0.013	0.063	195,442
Log(Market cap)	20.81	1.53	18.49	19.61	20.58	21.75	25.14	195,441
B/M	0.622	0.969	0.039	0.286	0.485	0.762	2.54	186866
S&P 500	0.172	0.377	0	0	0	0	1	195,490
Index %	2.656	2.165	0	0.553	2.362	4.559	6.961	195,490
MF %	21.83	13.64	0	11.27	21.07	30.93	56.46	195,490
Ins.%	44.45	18.87	0	32.32	45.43	57.28	85.46	195,490
N ETF holders	27.60	21.15	0	11	23	41	85	195,490
Abs(FIPP)	0.098	1.481	0.001	0.019	0.042	0.075	0.469	184,736

Table 2

Fund-level tests. Panel A presents fund-month panel regressions that relate three measures of ETF activity to underlying stock return correlations. *Fratio* is the ratio of the variance of the portfolio to the average of the variances of the stocks in the portfolio. *Holdings %* is the proportion of the portfolio's total market capitalization that is owned by the ETF on the last day of the prior month. *SD shares* is the standard deviation of ETF shares outstanding. *ETF turnover* is the average daily turnover of ETF shares.

Additional control variables and fixed effects appear in Panel B. *Expense ratio* is the annual expense ratio of the fund, in percent. *TNA* is total net assets of the fund as of the latest report, in millions of dollars. *Log(N holdings)* is the log of the number of the ETF's holdings of common stock that are also in our CRSP stock sample. Detailed variable definitions appear in the Appendix.

We follow Driscoll and Kraay (1998) to compute a nonparametric covariance matrix estimator that produces heteroskedasticity- and autocorrelation-consistent standard errors that are robust to general forms of spatial and temporal dependence.***, ** and * signify statistical significance at the 1%, 5% and 10% levels.

Panel A

	(1)	(2)	(3)	(4)
		Y = Fratio		
Holdings %	0.0800***			0.0654***
	(0.00682)			(0.00684)
SD shares		0.252***		0.130*
		(0.0641)		(0.0678)
ETF turnover			0.225***	0.184***
			(0.0328)	(0.0366)
Constant	0.448***	0.448***	0.445***	0.437***
	(0.0237)	(0.0230)	(0.0235)	(0.0236)
Time FE	NO	NO	NO	NO
Fund FE	NO	NO	NO	NO
Observations	24,799	24,799	24,799	24,799
R-squared	0.009	0.005	0.017	0.024

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Y = Fratio					
Holdings %	0.0466***			0.0438***	0.0186			0.0212
	(0.00841)			(0.00823)	(0.0136)			(0.0137)
SD shares		0.128***		-0.0294		0.0114		-0.0222
		(0.0342)		(0.0296)		(0.0140)		(0.0150)
ETF turnover			0.187***	0.189***			0.107***	0.113***
			(0.0178)	(0.0170)			(0.0281)	(0.0298)
Expense ratio	-4.956***	-5.008***	-6.422***	-6.455***	-1.344	-1.332	-1.013	-1.021
	(1.147)	(1.195)	(1.444)	(1.408)	(1.296)	(1.296)	(1.317)	(1.320)
Log(TNA)	-3.84e-05	0.00312***	-0.000213	-0.00326***	-0.00437*	-0.00369	-0.00360	-0.00440*
	(0.00108)	(0.000840)	(0.000805)	(0.00108)	(0.00235)	(0.00226)	(0.00219)	(0.00229)
Log(Nholdings)	-0.0489***	-0.0500***	-0.0500***	-0.0487***	-0.0535***	-0.0535***	-0.0533***	-0.0532***
	(0.00268)	(0.00280)	(0.00283)	(0.00266)	(0.00599)	(0.00599)	(0.00599)	(0.00598)
Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Fund FE	NO	NO	NO	NO	YES	YES	YES	YES
Observations	20,884	20,884	20,884	20,884	20,884	20,884	20,884	20,884
R^2	0.615	0.614	0.623	0.625	0.748	0.748	0.7493	0.7494

Table 3

Fund-level tests: Subsets. Coefficients from multivariate regressions on monthly data from 2006-2012. The dependent variable is $Fratio$, the ratio of the variance of the portfolio to the average of the variances of the stocks in the portfolio. *Holdings %* is the proportion of the portfolio's total market capitalizationthat is owned by the ETF on the last day of the prior month. *SD shares* is the standard deviation of ETF shares outstanding. *ETF turnover* is the average daily turnover of ETF shares. Detailed variable definitions appear in the Appendix. All regressions include all control variables in Table 2, Panel B and time and stock fixed effects. Panel A breaks the sample into stock market capitalization terciles where tercile 1 is the smallest. Panel B breaks the sample into terciles by R^2. Panel C presents subsets by type of holdings as described by the CRSP mutual fund database. Column (1) presents the full sample, column (2) presents the subsample of ETFs that hold corporate and municipal bonds (sample size 1,168), and column (3) presents the stocks that hold more than the sample median proportion of common stock. *abs(gap)* is the absolute value of the difference between the fund's reported net asset value and it's closing market price. Driscoll and Kraay (1998) standard errors are used, and ***, ** and * signify statistical significance at the 1%, 5% and 10% levels.

Panel A

$Y = Fratio$	TNA 1	TNA 2	TNA 3
Holdings %	-0.228	-0.0457	0.0191
	(0.176)	(0.0367)	(0.0140)
SD shares	-0.0180	-0.00221	0.0797***
	(0.0205)	(0.0224)	(0.0284)
ETF turnover	0.00427	0.0824**	0.137***
	(0.0211)	(0.0314)	(0.0301)

Panel B

Y = $Fratio$	R^2 1	R^2 2	R^2 3
Holdings %	0.00533	0.0114	0.0260
	(0.0275)	(0.0213)	(0.0198)
SD shares	0.0181	0.0104	0.0272
	(0.0296)	(0.0199)	(0.0290)
ETF turnover	0.0763	0.0805***	0.115***
	(0.0460)	(0.0253)	(0.0296)

Panel C

	(1) Full sample	(2) Illiquid>0	(3) Common $\geq$ median
Holdings %	0.0212	0.0570	-0.0108
	(0.0137)	(0.0684)	(0.0171)
SD shares	-0.0222	-0.128	-0.0154
	(0.0150)	(0.0843)	(0.0256)
ETF turnover	0.113***	0.105	0.120***
	(0.0298)	(0.104)	(0.0344)
Avg abs(gap)	0.0025	0.0056	0.0020

Table 4

Stock-level tests. Month/stock panel regressions relate three measures of a stock's exposure to ETF activities to two measures of its comovement with the market portfolio. β_M is the coefficient of the stock's daily excess returns on daily market excess returns in that month. *Sratio* is the ratio of stock variance to stock idiosyncratic variance of daily returns over the month. *ETF %* is the proportion of the stock that is held by ETFs, using CRSP data, on the last day of the prior month. *Wtd SD* is the weighted average percentage standard deviation in the shares outstanding of the ETFs that hold the stock. *Wtd turnover* is the weighted average turnover of the ETFs that hold the stock.

Additional control variables and fixed effects appear in Panels B and C. *Stock turnover* is the average daily turnover of the stock over the month *Log(Mkt cap)* is the log of the firm's market capitalization. *B/M* is the book-to-market ratio. *S&P 500* is an indicator variable for whether the stock is currently in the S&P 500 index. *MF %* and *Ins.%* are is the proportion of the stock held by mutual funds and other institutions, in percent. *N ETF holders* is the number of ETFs that hold the stock. *Abs(FIPP)* is the absolute value of flow-induced price pressure in the current quarter. Detailed variable definitions appear in the Appendix.

We follow Driscoll and Kraay (1998) heteroskedasticity- and autocorrelation-consistent standard errors that are robust to general forms of spatial and temporal dependence appear in parentheses. ***, ** and * signify statistical significance at the 1%, 5% and 10% levels. R^2 excludes the explanatory power of fixed effects.

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Y = β_M				Y = *Sratio*		
ETF %	0.0515***			0.0529***	0.219***			0.209***
	(0.00883)			(0.00925)	(0.0405)			(0.0395)
Wtd SD		2.414***		0.889		0.469		-6.331***
		(0.555)		(0.876)		(1.602)		(1.782)
Wtd turn			0.787***	0.656***			1.916***	2.839***
			(0.150)	(0.216)			(0.591)	(0.556)
Constant	1.059***	1.109***	1.083***	0.945***	2.063***	2.578***	2.349***	1.923***
	(0.0424)	(0.0302)	(0.0349)	(0.0495)	(0.0723)	(0.161)	(0.148)	(0.0669)
Time FE	NO	NO	NO	NO	NO	NO	NO	NO
Stock FE	NO	NO	NO	NO	NO	NO	NO	NO
N	195,490	195,490	195,490	195,490	195,490	195,490	195,490	195,490
R^2	0.018	0.006	0.014	0.032	0.057	0.000	0.014	0.075

Panel B

	(1)	(2)	(3)	(4)
		$Y = \beta_M$		
ETF %	0.0270***			0.0190***
	(0.00752)			(0.00619)
Wtd SD		1.357***		-0.522
		(0.427)		(0.398)
Wtd turn			0.686***	0.720***
			(0.145)	(0.134)
Stock turn	3.647***	3.640***	3.483***	3.447***
	(0.763)	(0.772)	(0.768)	(0.767)
Log(Mkt cap)	-0.186***	-0.181***	-0.176***	-0.179***
	(0.0348)	(0.0347)	(0.0348)	(0.0349)
B/M	0.00684	0.00909	0.0118	0.0108
	(0.00918)	(0.00918)	(0.00908)	(0.00907)
S&P 500	0.0341	-0.00725	-0.0499**	-0.0268
	(0.0243)	(0.0241)	(0.0235)	(0.0241)
Index %	-0.0277***	-0.0133	-0.00932	-0.0185*
	(0.00989)	(0.0106)	(0.0105)	(0.00976)
MF %	0.00584***	0.00511***	0.00483***	0.00526***
	(0.000785)	(0.000744)	(0.000722)	(0.000756)
Ins. %	0.00313***	0.00283***	0.00244***	0.00259***
	(0.000650)	(0.000671)	(0.000704)	(0.000697)
Abs(FIPP)	-0.00118	-0.00116	-0.000898	-0.000856
	(0.00181)	(0.00178)	(0.00180)	(0.00182)
Time FE	YES	YES	YES	YES
Stock FE	YES	YES	YES	YES
N	180,263	180,263	180,263	180,263
R^2	0.0651	0.0648	0.0696	0.0706

Panel C

	(1)	(2)	(3)	(4)
		$Y = Sratio$		
ETF %	0.0391***			-0.367
	(0.0136)			(1.212)
Wtd SD		1.748**		0.807***
		(0.692)		(0.284)
Wtd turn			0.816***	0.824***
			(0.183)	(0.283)
Stock turn	-12.44***	-12.44***	-12.61***	-12.68***
	(2.322)	(2.299)	(2.332)	(2.355)
Log(Mkt cap)	0.0325	0.0401	0.0452	0.0407
	(0.0331)	(0.0335)	(0.0335)	(0.0332)
B/M	-0.0212	-0.0181	-0.0151	-0.0166
	(0.0181)	(0.0184)	(0.0178)	(0.0176)
S&P 500	0.298***	0.240***	0.190***	0.228***
	(0.0461)	(0.0466)	(0.0525)	(0.0530)
Index %	0.00254	0.0230	0.0275	0.0132
	(0.0175)	(0.0187)	(0.0190)	(0.0184)
MF %	0.00605***	0.00502***	0.00470***	0.00537***
	(0.00124)	(0.00124)	(0.00127)	(0.00125)
Ins. %	-0.000150	-0.000563	-0.00102	-0.000776
	(0.00109)	(0.00111)	(0.00111)	(0.00110)
Abs(FIPP)	0.00107	0.00107	0.00138	0.00145
	(0.00320)	(0.00318)	(0.00319)	(0.00322)
Time FE	YES	YES	YES	YES
Stock FE	YES	YES	YES	YES
N	180,263	180,263	180,263	180,263
R^2	0.2951	0.2949	0.2959	0.2962

Table 5

Stock-level tests: Subsets by size, turnover and ETF holdings. Coefficients from stock-month panel regressions of measures of a stock's comovement with the market portfolio on measures of the activity of ETFs holding the stock. The sample period is July 2006 to June 2012. β_M is the coefficient of the stock's daily excess returns on daily market excess returns in that month. *Sratio* is the ratio of stock variance to stock idiosyncratic variance of daily returns over the month. *ETF %* is the proportion of the stock that is held by ETFs, using CRSP data, on the last day of the prior month. *Wtd SD* is the weighted average percentage standard deviation in the shares outstanding of the ETFs that hold the stock. *Wtd turnover* is the weighted average turnover of the ETFs that hold the stock. All regressions also contain all control variables in Table 4 (*Stock turnover*, *Log(Market cap)*, *B/M*, *S&P 500*, *MF%*, *Ins%* and *Abs(FIPP)*) and time and stock fixed effects. Detailed variable definitions appear in the Appendix. Panel A breaks the sample into stock market capitalization terciles where tercile 1 is the smallest. Panel B breaks the sample into terciles by turnover, which is calculated as volume divided by shares outstanding from CRSP. Driscoll and Kraay (1998) heteroskedasticity- and autocorrelation-consistent standard errors that are robust to general forms of spatial and temporal dependence appear in parentheses. ***, ** and * signify statistical significance at the 1%, 5% and 10% levels. R^2 excludes the explanatory power of fixed effects.

Panel A

	Size 1	Size 2	Size 3
Y = β_M			
ETF %	0.0763***	0.0212***	0.00905**
	(0.0123)	(0.00587)	(0.00404)
Wtd SD	2.018***	0.682**	1.035*
	(0.574)	(0.342)	(0.556)
Wtd turnover	1.193***	0.277***	0.448***
	(0.162)	(0.0852)	(0.159)
Y = $Sratio$			
ETF %	0.108***	0.0310***	0.0281***
	(0.0145)	(0.00993)	(0.0106)
Wtd SD	1.768***	0.479	1.688*
	(0.596)	(0.657)	(0.903)
Wtd turnover	1.087***	0.392***	1.105***
	(0.169)	(0.0980)	(0.283)

Panel B

	Turn 1	Turn 2	Turn 3
Y = β_M			
ETF %	0.0437***	0.0101**	0.0183***
	(0.0151)	(0.00483)	(0.00681)
Wtd SD	2.338***	0.903**	0.478
	(0.632)	(0.410)	(0.441)
Wtd turnover	1.292***	0.327***	0.298***
	(0.190)	(0.109)	(0.111)
Y = $Sratio$			
ETF %	0.0997***	0.0281**	0.0114
	(0.0306)	(0.0117)	(0.0109)
Wtd SD	2.721***	2.413**	1.208
	(0.782)	(1.075)	(0.729)
Wtd turnover	1.220***	0.675**	0.572***
	(0.248)	(0.261)	(0.143)

Table 6

Correlation between two largest ETF stock holdings regressed on weighted FIPP of remaining holdings. The monthly correlation of daily returns of the two largest holdings of each ETF is regressed on the absolute value of the weighted-average FIPP of the remaining holdings, *Abs(Wtd* $FIPP_{\{1,2\}^{\perp}}$*)*. Control variables are the absolute value of the FIPP of each of the two largest holdings, *Abs(*$FIPP_1$*)* and *Abs(*$FIPP_2$*)*, the log of the market capitalization of each holding, *Log(*$Marketcap_1$*)* and *Log(*$Marketcap_2$*)*, the percentage of the fund's TNA of each holding, *Percent* TNA_1 and *Percent* TNA_2, the fund's expense ratio, *Expense ratio*, the fund's TNA, *Log*(TNA), the number of holdings of the fund, *Log*($Nholdings$), and fund and month fixed effects. Driscoll and Kraay (1998) heteroskedasticity- and autocorrelation-consistent standard errors that are robust to general forms of spatial and temporal dependence appear in parentheses. ***, ** and * signify statistical significance at the 1%, 5% and 10% levels. R^2 excludes the explanatory power of fixed effects.

	(1)	(2)	(3)	(4)
		Y = $Corr(R_1, R_2)$		
Abs(Wtd $FIPP_{\{1,2\}^{\perp}}$)	0.0801***	0.0912***	0.0881***	0.103***
	(0.0218)	(0.0266)	(0.0243)	(0.0265)
Abs($FIPP_1$)			0.00969	0.00540
			(0.0153)	(0.0150)
Abs($FIPP_2$)			0.0125***	0.0122***
			(0.00417)	(0.00459)
Log(Marketcap_1)			0.00576	0.00348
			(0.00373)	(0.00403)
Log(Marketcap_2)			-0.00924	-0.0139
			(0.0157)	(0.0165)
Percent TNA_1			0.00309	-0.167*
			(0.0207)	(0.0935)
Percent TNA_2			0.0604*	0.383***
			(0.0312)	(0.124)
Expense ratio		-7.637**		-8.234**
		(3.428)		(3.462)
Log(TNA)		-0.00114		-0.000878
		(0.00430)		(0.00476)
Log(Nholdings)		0.0230*		0.00121
		(0.0129)		(0.0112)
Time FE	YES	YES	YES	YES
Fund FE	YES	YES	YES	YES
Observations	24,760	20,872	22,467	19,752
R-squared	0.2279	0.232	0.253	0.2528

Table 7

Mean betas on lagged market returns by tercile of ETF activity. Betas come from the monthly stock-by-stock regression $R_t = \beta R_t^M + \beta_{R_{t-1}^M} R_{t-1}^M + \beta_{R_{t-2}^M} R_{t-2}^M + \beta_{R_{t-3}^M} R_{t-3}^M + \beta_{R_{t-4}^M} R_{t-4}^M + \epsilon_t$ where R_t^M is the return on the market portfolio and R_t is the daily return on the stock. Tercile 1 is the tercile with the lowest ETF activity, and Tercile 3 is the tercile with the highest ETF activity. N = 72 year-month observations.

	$\beta_{R_{t-1}^M}$	$\beta_{R_{t-2}^M}$
Terciles by ETF %		
1	0.0272**	0.0061
	(0.013)	(0.009)
3	-0.0216*	-0.0029
	(0.013)	(0.016)
Difference	0.0488***	0.0089
	(0.010)	(0.008)
Terciles by Wtd SD		
1	0.0489***	-0.0008
	(0.008)	(0.007)
3	-0.0376**	-0.0038
	(0.016)	(0.014)
Difference	0.0864***	0.0030
	(0.012)	(0.011)
Terciles by Wtd turn		
1	0.0656***	0.0054
	(0.009)	(0.007)
3	-0.0579**	-0.0071
	(0.016)	(0.014)
Difference	0.1235***	0.0125
	(0.012)	(0.010)

Table 8

Lag market betas regressed on measures of ETF activity. Betas come from the monthly stock-by-stock regression $R_t = \beta R_t^M + \beta_{R_{t-1}^M} R_{t-1}^M + \beta_{R_{t-2}^M} R_{t-2}^M + \beta_{R_{t-3}^M} R_{t-3}^M + \beta_{R_{t-4}^M} R_{t-4}^M + \epsilon_t$ where R_t^M is the return on the market portfolio and R_t is the daily return on the stock. *ETF %* is the proportion of the stock that is held by ETFs, using CRSP data, on the last day of the prior month. *Wtd SD* is the weighted average percentage standard deviation in the shares outstanding of the ETFs that hold the stock. *Wtd turnover* is the weighted average turnover of the ETFs that hold the stock. *Stock turnover* is the average daily turnover of the stock over the month *Log(Mkt cap)* is the log of the firm's market capitalization. *B/M* is the book-to-market ratio. *S&P 500* is an indicator variable for whether the stock is currently in the S&P 500 index. *MF %* and *Ins.%* are is the proportion of the stock held by mutual funds and other institutions, in percent. *N ETF holders* is the number of ETFs that hold the stock. *Abs(FIPP)* is the absolute value of flow-induced price-pressure. Detailed variable definitions appear in the Appendix. Driscoll and Kraay (1998) heteroskedasticity- and autocorrelation-consistent standard errors that are robust to general forms of spatial and temporal dependence appear in parentheses. ***, ** and * signify statistical significance at the 1%, 5% and 10% levels. R^2 excludes the explanatory power of fixed effects.

	(1)	(2)
	$\beta_{R^M_{t-1}}$	$\beta_{R^M_{t-2}}$
ETF %	-0.00444*	0.000779
	(0.00243)	(0.00312)
Wtd SD	-0.0459	-0.264
	(0.404)	(0.188)
Wtd turn	-0.377***	-0.114**
	(0.0737)	(0.0441)
Stock turn	4.354***	0.598
	(0.868)	(0.887)
Log(Mkt cap)	-0.0401*	-0.0320*
	(0.0218)	(0.0188)
B/M	0.00971	0.00676*
	(0.00808)	(0.00378)
S&P 500	-0.0194	0.0153
	(0.0226)	(0.0232)
Index %	0.0165***	0.00761*
	(0.00428)	(0.00424)
MF %	-0.000531	-0.000128
	(0.000505)	(0.000410)
Ins. %	-0.000925**	0.000112
	(0.000368)	(0.000264)
Abs(FIPP)	-0.000778	-0.000158
	(0.000535)	(0.000818)
Time FE	YES	YES
Stock FE	YES	YES
N	180,263	180,263
R^2	0.0299	0.0221