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

Zhi Da and Sophie Shive*

First Version: Aug 2012
This Version: Mar 2013

Abstract

We provide novel empirical evidence that arbitrageurs transmit non-fundamental shocks from exchange-traded funds (ETFs) to the broad cross-section of stocks they hold, thus making these shocks systematic. Using a large sample of US equity ETFs from 2006 to 2012, we find a strong relation between measures of ETF activity and return comovement at both the fund level and the stock level, after controlling for time and stock fixed effects and a host of other controls including institutional trading. ETF trading activity increases a stock's exposure to sentiment risk and makes its loading on lagged market return more negative, suggesting that at least some return comovement is excessive. The effect of ETF activity is stronger among small and illiquid stocks and during periods of turbulent markets. The link between ETF activity and return comovement weakens immediately following the recent flash crash as many high frequency traders who are natural ETF arbitrageurs chose to stay on the sidelines.

*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 and seminar at University of Notre Dame, Malcom Baker, Robert Battalio, Martijn Cremers, Ben Golez, Robin Greenwood, Paul Schultz, Mao Ye, and Xiaoyan Zhang. Errors are ours.

1 Introduction

Perhaps due to a half-century of encouragement from finance academics, investment assets are increasingly indexed. 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 have made indexing cheaper and more convenient for many investors. Exchange traded funds had $1.2 trillion in assets under management in early 2012.[1] According to the ETF industry association, net cash flows into exchange traded funds have been $38,650, $55,115, and $72,260 million in 2009, 2010 and 2011.[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 push stock prices by simultaneously taking opposite positions in the ETF and the underlying basket 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 on to their component stocks. Arbitrageurs, who are generally enforcers of price efficiency, can thus at times make "mis-pricing" more systematic, consistent with results in Shleifer and Vishny (1997), Hong, Kubik, and Fishman (2012) and Lou and Polk (2013).

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

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

[2] http://www.etf-ia.com/sites/all/themes/etfia/downloads/June-2012-ETF-Data-Net-Flows.pdf

component stocks move together, controlling for fund and time fixed effects, a host of fund-level control variables, and correcting for cross-correlation in error terms. An 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.

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. Creation and redemption is more onerous than trading and typically occurs after market close, when an arbitrage opportunity may no longer be available.[3]

We then focus our attention on stock-level analysis. 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, 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 0.078 increase in a stock's CAPM beta, controlling for stock and time fixed effects and a host of stock-level control variables. We find weaker 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. These results hold with respect to different measures of stock's comovement with the market. Finally, the effect of ETF activities on stock comovement is stronger among small stocks, stocks with low

[3]See Abner (2010).

turnover, and stocks with higher ETF holdings.

A natural question follows after we establish a positive link between ETF activities and return comovement: 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 by first looking at lagged market betas. Empirically, we find that stocks with higher measures of ETF activity tend to have negative betas on lagged market returns, and that a stock's 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 reversals in prices, a symptom of "excess" comovement.

We also examine the change in investor sentiment as a specific non-fundamental shock. We compute the sentiment beta of a stock as the regression coefficient of the stock's monthly stock excess return on the changes in the investor sentiment index of Baker and Wurgler (2006). We find that a one standard deviation increase in the weighted average ETF turnover measure is associated with a 0.006 increase in a stock's sentiment beta, more than half its mean and almost 10% of its standard deviation. ETF trading appears to make its underlying stocks more prone to sentiment risk, likely via the arbitrage trading channel.

Finally, we present two pieces of time-series evidence that further link ETF arbitrage activities to return comovement. First, if excessive return comovement comes from non-fundamental shocks that are made systematic by ETF arbitrage, we would expect our results to be stronger during periods of turbulent market when such shocks are plenty that drive the initial price wedge between the ETF and its underlying stocks. 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 the arbitrage opportunity measured by the profitability of the high frequency traders who are the natural ETF arbitrageurs.[4] When we correlate lagged monthly VIX index to our

[4] 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

monthly cross-sectional regression coefficient that identifies the marginal impact of ETF trading on return comovement, we find a correlation of more than 50%. Such a high correlation confirms that ETF trading has a bigger impact on return comovement precisely during periods when ample arbitrage opportunities are available.

We also take advantage of the flash crash on May 6, 2010 as a relatively exogenous negative shock to arbitrage activity. We find that the link between ETF activity and return comovement weakens immediately following the flash crash as high-frequency traders who are natural ETF arbitrageurs chose to stay on the sideline.[5] Overall, our paper provides novel empirical evidence that arbitrageurs can at times exacerbate stock comovement and mispricing risk, thus highlighting an important form of limits to arbitrage.

Our paper is related to the extensive literature on the drivers of equity return correlation. This 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 (Greenwood and Sosner (2007) for the Nikkei 225, Goetzmann and Massa (2003) and Barberis, Shleifer, and Wurgler (2005) for the S&P 500). This literature is subject to the possibility 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) and Anton and Polk (2013). To the extent that institutions have

trading firms, he said. Such firms' profits from equities trading have fallen 75%, from 2009."

[5] According to a WSJ on the anniversary of the flash crash, "companies that use fast-trading, computer-driven strategies, which were painted by some as culprits of the collapse, have curtailed trading."

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

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. Another stream of this literature focuses on the drivers of differences between the market price of the exchange-traded fund and the price of the underlying portfolio. In an efficient market, the two prices should be close accounting for trading frictions. Engle and Sarkar (2002) find that the price discrepancies for domestic ETFs are generally small and short-lived. Marshall, Nguyen, and Visaltanachoti (2012) find that they quickly correct and are accompanied by a decrease in liquidity. Petajisto (2011) finds that deviations are larger in funds holding international or illiquid securities where net asset values are most difficult to compute. Jiang and Yan (2012) 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) finds that arbitrage activity propagates liquidity shocks from ETFs to the underlying stocks and increases volatility, but does not investigate stock comovement and mispricing risk. 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 fund level and 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. With the growth and popularity

[7] see www.etfdb.com.

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 ETFs and other index trading activities coincides with rising return correlations among stocks. Taking advantage of a large panel of ETF holdings, we are able to better quantify the incremental impact of ETFs on equity comovement and mispricing risk and our findings are clearly relevant for developing the ETFs market going forward.

The paper proceeds as follows. The following section presents the data used in the study, Section 3 presents the empirical link between various ETF activities and return comovement at both the fund- and stock-level, Section 4 confirms that at least part of such return comovement is excessive, Section 5 presents time-series evidence, and the last section concludes.

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

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

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 portfolio. The median fund holds 0.012% of its underlying portfolio. Thus, 678 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 5% 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 5% 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 is 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 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 at both the fund-level and 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. To gain 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 and median close to 0.46.

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. Note that, for simplicity, this is an overall measure and does not vary according to the weights in each individual stock that the ETF holds.

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 2.79% of ETF shares outstanding per day. The median is smaller at 1.02%.

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 5% per day.

We will use fund and time fixed effects, which will 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

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

time and fund fixed effects and control variables. Columns 5-8 have only time fixed effects and columns 9-12 have both time and fund fixed effects. When only time fixed effects are included, *Holdings%* and *SD shares* are still significant (columns 5-6), but they are no longer significant when fund fixed effects are added. In contrast, *ETF turnover* is significant in all regression specifications.

In column 12, 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 a little less than 2.6% of its mean value, and 14.5% of its standard deviation.

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 itself 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.[10]

3.2 Stock-level tests

From this point onwards, 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 comove more with the market. In addition, if the initial ETF price change reflects a non-fundamental liquidity or sentiment shock, then subsequent arbitrage trading in stocks will push the stock prices in the same direction as this shock. Thus, we might expect a stock with greater exposures to ETF activities to be more prone to sentiment risk. We test these predictions using stock-level data.

3.2.1 Empirical measures

In order to perform 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 three measures of how much a stock comoves with the market. The first, in the spirit of Pollet and Wilson (2008), is the total variance of the stock divided by the idiosyncratic variance. This number will be high if the stock 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,

[10] 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 mis-pricing. Second and more importantly, a price difference can reflect either an actual opportunity for arbitrage trade or the presence of limits-to-arbitrage.

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. As with the $Fratio$, we winsorize $Sratio$ at the 1% level to remove the effect of outliers.

A second, more conventional measure of how much a stock moves with the market is it's one-factor CAPM beta, β_M. This is the beta obtained from a regression of excess stock returns on the excess market returns provided by Kenneth French's Web site. A third measure is the square root of the R^2 of the same regression. We call this variable ρ_M since it is equivalent to the correlation coefficient between the stock excess return and the market excess return in a univariate CAPM regression. We also use the Fama and French (1993) and Carhart (1997) 4-factor model, but results are similar and untabulated. All three of these variables are winsorized at the 1% level to mitigate the effect of outliers. Summary statistics of these variables appear in Table 1, Panel B.

We have three measures of ETF activity at the stock level. The first measure, *ETF%*, is the 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*1000). We also include the log of the stock's market capitalization from CRSP and its Book/Market 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 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

FIPP variable in a shorter sample 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 all three 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 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 through D present multivariate results when both time and stock fixed effects are included. All coefficients on the three explanatory variables of interest are lower. For example, in Panel C, β_M is the dependent variable. Column 5 shows that the coefficient on ETF holdings is significant controlling for date and fund fixed effects and control variables. Importantly, the effect of ETF holdings is more than three times the effect of mutual fund holdings or of institutional holdings. In this column, a 1% increase in ETF holdings of a stock is associated with a 0.0295 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.

[11] We thank Dong Lou for supplying us with his FIPP data. Since the data ends in 2009, we do not include it in our main analysis.

Wtd SD is significant on its own with fixed effects and controls (see column 6), but when it is with the other two explanatory variables of interest, its sign changes (column 8). 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 3, Panel C, column 7 shows that a one standard deviation increase in *Wtd turnover* is associated with a 0.693*0.112=0.0776 increase in β_M. This effect hardly changes when *Wtd turnover* is with the other two explanatory variables of interest.

The effects are similar with *Sratio* and ρ_M. 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. Panel D, Column 9 shows that 1% increase in ETF holdings is associated with an approximate 0.972% increase in its correlation with the market (ρ_M). Columns 11 shows that a one standard deviation increase in *Wtd turnover* is associated with an 0.200*0.112=2.24% increase in ρ_M.

3.2.3 Sub-sample cuts

Table 5 shows Table 4's results broken down by terciles of size, turnover, and ETF holdings. 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 generally not in the others. For smaller and lower turnover stocks, it must be more difficult to locate or sell components during creation and redemption. Panel C

breaks down the sample by ETF holdings, and shows that the effect is also strongest for stocks with higher ETF holdings where ETF-related arbitrages have relatively bigger effect on the turnover of the stock.

4 Is the return co-movment excessive?

So far, we have provided evidence suggesting that ETF activity is positively related to return comovements. 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 Lagged betas

We first 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 load negatively on the market return on day t. In other words, the stock will have negative lagged market betas. If higher return comovement comes from faster incorporation of common information, we should not observe return reversals and hence negative lagged betas.

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 6 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 6 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 7, 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 6, the results are strongest for *Wtd turn*.

4.2 Sentiment beta

To further examine whether higher return comovement induced by ETF trading reflects non-fundamental shocks, we examine one such shock, change in investor sentiment. Specifically, we run similar regressions with the sentiment beta, measuring a stock's exposure to sentiment risk, as the dependent variable.

We compute sentiment beta (β_{SENT}) with the following regression for each year and stock i:

$$R_{i,t} - R_{f,t} = \beta_i(R_{M,t} - R_{f,t}) + \beta_{SENT,i}\Delta SENT_t^{\perp} + \epsilon_{i,t}, \tag{10}$$

where $R_{i,t} - R_{f,t}$ is the monthly excess return on stock i, $R_{M,t} - R_{f,t}$ is the monthly market excess return, and $\Delta SENT^{\perp}$ is the orthogonalized change in monthly Baker and Wurgler (2006) investor sentiment index. Given the monthly frequency associated with the sentiment index, the regression is run every year for each stock with at least 8 months of data. Table 1 Panel B suggests

that β_{SENT} has a mean close to zero in our sample and its standard deviation, 0.074, is also much smaller compared to that of the CAPM beta (0.747).

Since the sentiment beta is estimated annually and the sentiment data ends in 2010, the time series length of the panel data is only five. As such, we include year fixed effects but not stock fixed effects in the regressions. We continue to use Driscoll and Kraay (1998) standard errors which are robust to heteroskedasticity and auto-correlated and cross-correlated error terms.

The results in Table 8 are consistent with our conjecture that arbitrage activities between an Exchange-traded fund (ETF) and its underlying stocks could propagate non-fundamental shocks from the ETFs to a broad cross-section of stocks they hold. We find that ETF trading activity increases a stock's exposure to sentiment risk. Column 3 shows that an one standard deviation increase in *Wtd turnover* is associated with a 0.0548*0.112=0.006 increase in β_{SENT}. Table 9 breaks these results down by tercile of size (Panel A) and stock turnover (Panel B). This table shows that the effect of ETFs on stocks is strongest for smaller and lower turnover stocks, precisely the set of stocks that are prone to sentiment shocks.

5 Time-series evidence

In this section, we present two pieces of time-series evidence that further link arbitrage activities to return comovement.

First, 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. Each month, we run cross-sectional regressions of a measure of stock comovement with the market ($Sratio$, β_M, or ρ_M) on weighted average turnover of the ETFs ($Wtd.turnover$) and other controls. The regression coefficient on $Wtd.turnover$ identify the marginal impact of ETF trading on return comovement in that month. Figure 2 plots the lagged VIX index against these regression coefficients. We observe a positive and significant correlation

coefficient regardless which comovement measure is used (corr = 0.55 for $Sratio$; corr = 0.21 for β_M; corr = 0.35 for ρ_M). These high correlations provide some support that ETF trading has a bigger impact on return comovement precisely during periods with ample mispricings.

We also take advantage of the flash crash on May 6, 2010 as a relatively exogenous negative shock to the arbitrage activity. Table 10 presents the coefficients of the three measures of stock comovement with the market on their *Wtd. turn* around the May 6, 2010 flash crash. the 6 months before is the "pre" period and 6 months after the crash is the "post period. We find that the link between ETF activity and return comovement weakens in the 6 months following the flash crash as high-frequency traders who are natural ETF arbitrageurs chose to stay on the sideline.

Overall, these preliminary time-series evidence provides further support that arbitrageurs can at times exacerbate stock comovement and mispricing risk, thus highlighting an important form of limits to arbitrage.

6 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 shares outstanding, meant to capture 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 an increased efficiency in reflecting market shocks. First, we find that sensitivity to lags of market returns is more negative stocks with high ETF ownership and decreases with measures of ETF activity, controlling for other possible drivers of this relationship. We also compute individual stock sentiment betas and find that they are higher for stocks that have higher ownership and turnover related to ETFs. This suggests that price changes that are not related to fundamentals can be propagated to stocks through the ETFs holding them. Finally, evidence from time series dimension provides further support that arbitrageurs can at times exacerbate stock return comovement and mispricing risk, thus highlighting an important form of limits to arbitrage.

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.

Baker, Malcolm, and Jeffery Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *The Journal of Finance* 61, 1645–1680.

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.

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.

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

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.

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.
	Stock-level measures
β_M	Coefficient of the stock's daily excess returns on daily market excess returns in that month.
β^{SENT}	Regression coefficient on $\Delta SENT^{\perp}$ from regressing monthly stock excess returns on monthly market excess returns and $\Delta SENT^{\perp}$ where $\Delta SENT^{\perp}$ denotes the orthogonalized change in monthly Baker and Wurgler (2006) sentiment index.
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 Thomson institution 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.
ρ_M	The square root of the R-squared from the regression of daily excess stock returns on excess market returns in that month.
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.

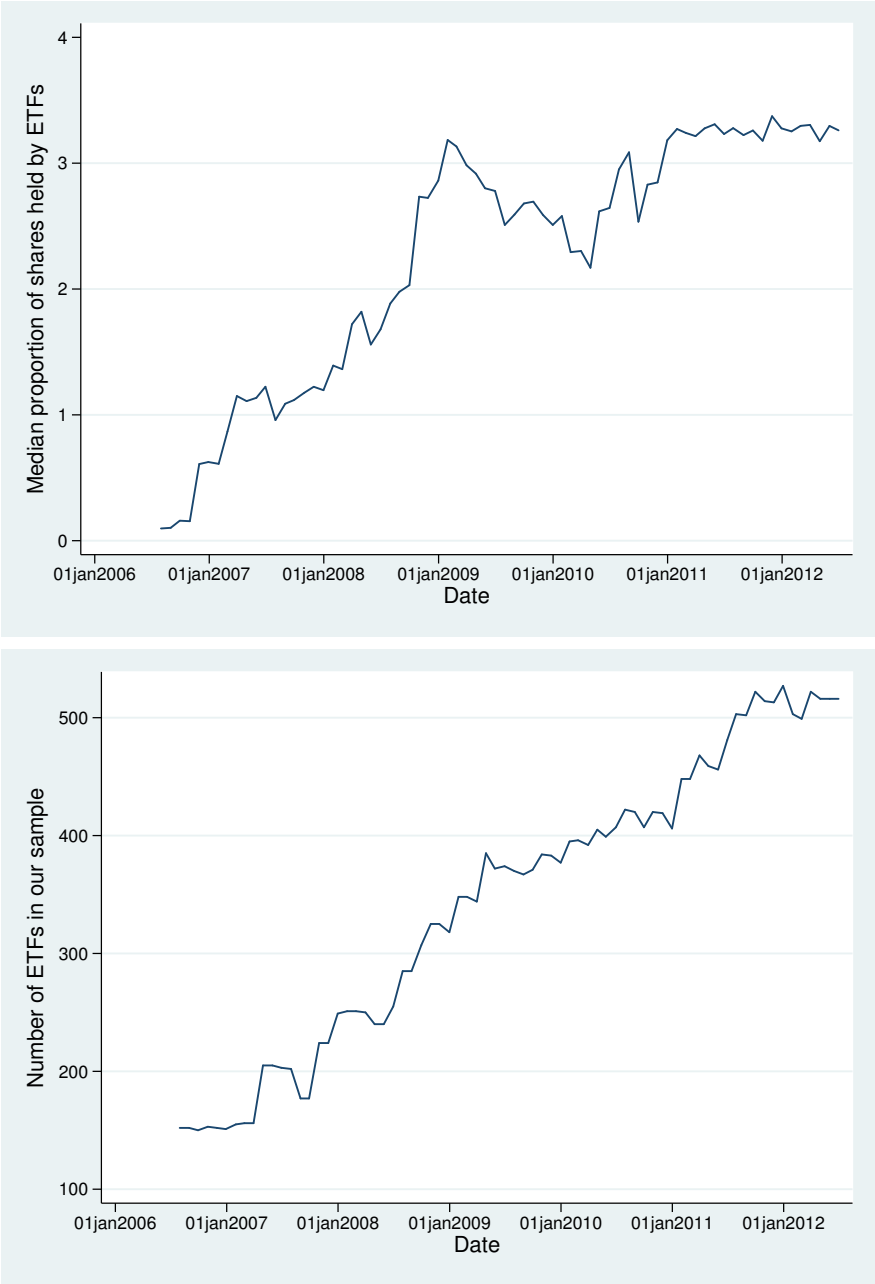

Figure 1. The top figure presents the median of the percentage of the stock that is held by exchange traded funds for CRSP stocks with share price of at least \$5 and market capitalization of at least \$100 million. The bottom figure presents the number of ETFs in our sample each month. The sample consists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

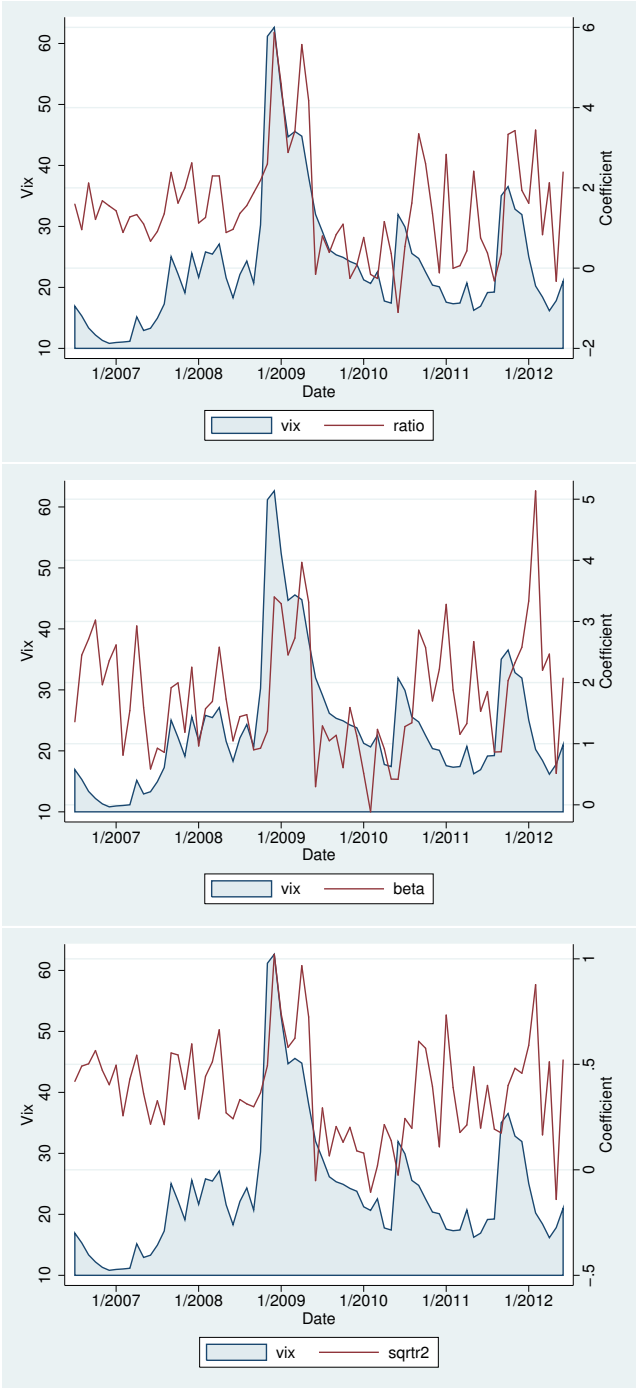

Figure 2. Coefficients of stock-level measures of correlation with the market, *Sratio* (top), β (middle) and $sqrt(R_M^2)$ (bottom) regressed on ETF weighted average turnover of those stocks, *Wtd turn*. One-month lagged VIX also appears in each figure.

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 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. *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.

Panel B presents stock-level summary statistics for 4,703 stocks. *S Ratio* 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. ρ_M is the R-squared from the regression of daily excess stock returns on excess market 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. Detailed variable definitions appear in the Appendix.

Panel A

Variable	Mean	SD	p1	p25	p50	p75	p99	N
FRatio	0.4555	0.1800	0.1341	0.3121	0.4380	0.5879	0.8681	24,799
Holdings %	0.0929	0.2173	0.0001	0.0023	0.0112	0.0584	1.2282	24,799
SD shares	0.0279	0.0483	0	0	0.0102	0.0328	0.2833	24,799
ETF turnover	0.0457	0.1046	0	0.0063	0.0119	0.0281	0.6550	24,799
Expense ratio	0.0050	0.0028	0.001	0.0028	0.0048	0.0065	0.0095	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

Panel B

S Ratio	2.592	1.789	1.036	1.457	1.987	3.018	11.129	195,490
β_M	1.183	0.755	-0.620	0.701	1.125	1.610	3.527	195,490
ρ_M	0.537	0.236	0.019	0.368	0.572	0.728	0.927	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	186,860
S&P 500	0.157	0.364	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

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 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.

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

	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
				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 by size and R^2. Coefficients from multivariate regressions on monthly data from 2006-2011. 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. 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 1	R 2	R 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)

Table 4

Stock-level tests. Month/stock panel regressions relate three measures of a stock's exposure to ETF activities to three measures of its comovement with the market portfolio. *S Ratio* 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. ρ_M is the R-squared from the regression of daily excess stock returns on excess market 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.

Additional control variables and fixed effects appear in Panels B, C and D. *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. 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)	(9)	(10)	(11)	(12)
			Y = $SRatio$				Y = β_M				Y = $sqrt(R_M^2)$	
ETF %	0.219***			0.209***	0.0515***			0.0529***	0.0360***			0.0343***
	(0.0405)			(0.0395)	(0.00883)			(0.00925)	(0.00334)			(0.00289)
Wtd SD		0.469		-6.331***		2.414***		0.889		0.130		-1.098***
		(1.602)		(1.782)		(0.555)		(0.876)		(0.256)		(0.247)
Wtd turn			1.916***	2.839***			0.787***	0.656***			0.341***	0.501***
			(0.591)	(0.556)			(0.150)	(0.216)			(0.0839)	(0.0707)
Constant	2.063***	2.578***	2.349***	1.923***	1.059***	1.109***	1.083***	0.945***	0.450***	0.533***	0.494***	0.425***
	(0.0723)	(0.161)	(0.148)	(0.0669)	(0.0424)	(0.0302)	(0.0349)	(0.0495)	(0.0139)	(0.0201)	(0.0205)	(0.0123)
Time FE	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Stock FE	NO	NO	NO	NO	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	195,490	195,490	195,490	195,490
R^2	0.057	0.000	0.014	0.075	0.018	0.006	0.014	0.032	0.088	0.000	0.026	0.121

Panel B.

	(1)	(2)	(3)	(4)
		$Y = Sratio$		
ETF %	0.0392***			0.0309***
	(0.0135)			(0.0116)
Wtd SD		1.712***		-0.331
		(0.646)		(1.099)
Wtd turn			0.778***	0.767***
			(0.175)	(0.261)
Stock turn	-5.503***	-5.508***	-5.574***	-5.592***
	(0.945)	(0.937)	(0.954)	(0.962)
Log(Mkt cap)	0.0256	0.0307	0.0340	0.0306
	(0.0352)	(0.0353)	(0.0354)	(0.0352)
B/M	-0.0258	-0.0234	-0.0210	-0.0223
	(0.0163)	(0.0164)	(0.0159)	(0.0158)
S&P 500	0.304*	0.279*	0.246	0.266
	(0.165)	(0.166)	(0.167)	(0.166)
Index %	0.000570	0.0204	0.0244	0.0102
	(0.0163)	(0.0169)	(0.0171)	(0.0170)
MF %	0.00594***	0.00513***	0.00495***	0.00549***
	(0.00107)	(0.00106)	(0.00109)	(0.00109)
Ins. %	-0.000450	-0.000760	-0.00112	-0.000941
	(0.000891)	(0.000904)	(0.000915)	(0.000906)
Time FE	YES	YES	YES	YES
Stock FE	YES	YES	YES	YES
N	186,860	186,860	186,860	186,860
R^2	0.3029	0.3021	0.3028	0.3036

Panel C

	(5)	(6)	(7)	(8)
		$Y = \beta_M$		
ETF %	0.0295***			0.0218***
	(0.00781)			(0.00645)
Wtd SD		1.484***		-0.396
		(0.422)		(0.382)
Wtd turn			0.693***	0.708***
			(0.138)	(0.125)
Stock turn	1.212***	1.204***	1.144***	1.132***
	(0.336)	(0.338)	(0.336)	(0.334)
Log(Mkt cap)	-0.188***	-0.184***	-0.181***	-0.184***
	(0.0352)	(0.0350)	(0.0351)	(0.0351)
B/M	0.00991	0.0119	0.0140	0.0131
	(0.00859)	(0.00854)	(0.00842)	(0.00843)
S&P 500	0.0871*	0.0688	0.0392	0.0525
	(0.0497)	(0.0541)	(0.0559)	(0.0534)
Index %	-0.0302***	-0.0150	-0.0114	-0.0215**
	(0.00968)	(0.0104)	(0.0101)	(0.00931)
MF %	0.00649***	0.00586***	0.00570***	0.00608***
	(0.000763)	(0.000713)	(0.000680)	(0.000717)
Ins. %	0.00381***	0.00356***	0.00323***	0.00336***
	(0.000595)	(0.000610)	(0.000630)	(0.000626)
Time FE	YES	YES	YES	YES
Stock FE	YES	YES	YES	YES
N	186,860	186,860	186,860	186,860
R^2	0.0587	0.0595	0.0639	0.0644

Panel D

	(9)	(10)	(11)	(12)
		$Y = \rho_M$		
ETF %	0.00972***			0.00746***
	(0.00168)			(0.00131)
Wtd SD		0.395***		-0.160
		(0.103)		(0.129)
Wtd turn			0.200***	0.209***
			(0.0337)	(0.0377)
Stock turn	-1.429***	-1.430***	-1.448***	-1.452***
	(0.110)	(0.109)	(0.113)	(0.114)
Log(Mkt cap)	0.00488	0.00613	0.00701	0.00620
	(0.00815)	(0.00808)	(0.00802)	(0.00805)
B/M	-0.00440***	-0.00382***	-0.00318**	-0.00350**
	(0.00139)	(0.00142)	(0.00142)	(0.00139)
S&P 500	0.0380**	0.0319**	0.0234	0.0278*
	(0.0151)	(0.0154)	(0.0157)	(0.0152)
Index %	-0.00570***	-0.000828	0.000276	-0.00319*
	(0.00188)	(0.00200)	(0.00190)	(0.00175)
MF %	0.00145***	0.00125***	0.00120***	0.00134***
	(0.000179)	(0.000166)	(0.000155)	(0.000165)
Ins. %	0.000915***	0.000840***	0.000743***	0.000786***
	(0.000108)	(0.000107)	(0.000106)	(0.000106)
Time FE	YES	YES	YES	YES
Stock FE	YES	YES	YES	YES
N	186,860	186,860	186,860	186,860
R^2	0.2624	0.2619	0.2655	0.2666

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 December 2011. *S Ratio* 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. ρ_M is the R-squared from the regression of daily excess stock returns on excess market 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. All regressions also contain all control variables in Table 4 (*Stock turnover, Log(Market cap), B/M, S&P 500, MF% and Ins%*) 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. Panel C breaks the sample into terciles by ETF holding %. 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

overall	Size 1	Size 2	Size 3
Y = $Sratio$			
ETF %	0.112***	0.0296***	0.0230**
	(0.0151)	(0.00882)	(0.00931)
Wtd SD	1.807***	0.363	0.982
	(0.612)	(0.656)	(0.852)
Wtd turnover	1.083***	0.383***	0.983***
	(0.166)	(0.0974)	(0.261)
Y = β_M			
ETF %	0.0794***	0.0224***	0.00960**
	(0.0126)	(0.00573)	(0.00399)
Wtd SD	2.043***	0.736**	0.986*
	(0.559)	(0.320)	(0.564)
Wtd turnover	1.156***	0.304***	0.435***
	(0.158)	(0.0754)	(0.159)
Y = $sqrt(R^2)$			
ETF %	0.0293***	0.00620***	0.00258**
	(0.00281)	(0.00132)	(0.00123)
Wtd SD	0.677***	0.0523	-0.0751
	(0.166)	(0.0887)	(0.128)
Wtd turnover	0.381***	0.0813***	0.0396
	(0.0414)	(0.0132)	(0.0386)

Panel B

	Turn 1	Turn 2	Turn 3
Y = $Sratio$			
ETF %	0.102***	0.0282**	0.00931
	(0.0313)	(0.0116)	(0.0101)
Wtd SD	2.487***	1.888*	1.191*
	(0.753)	(1.023)	(0.671)
Wtd turnover	1.104***	0.560**	0.544***
	(0.243)	(0.243)	(0.122)
Y = β_M			
ETF %	0.0439***	0.0118**	0.0196***
	(0.0155)	(0.00502)	(0.00658)
Wtd SD	2.325***	0.818*	0.578
	(0.592)	(0.413)	(0.426)
Wtd turnover	1.211***	0.311***	0.310***
	(0.177)	(0.107)	(0.110)
Y = $sqrt(R^2)$			
ETF %	0.0184***	0.00613***	0.00525***
	(0.00445)	(0.00143)	(0.00114)
Wtd SD	0.698***	0.246**	0.144*
	(0.166)	(0.108)	(0.0852)
Wtd turnover	0.368***	0.0999***	0.0744***
	(0.0502)	(0.0324)	(0.0219)

Panel C

	ETF% 1	ETF% 2	ETF% 3
Y = $Sratio$			
ETF %	0.237***	0.143***	0.00493
	(0.0380)	(0.0364)	(0.0159)
Wtd SD	1.028**	1.591	0.636
	(0.504)	(1.282)	(1.576)
Wtd turnover	0.306**	0.912***	0.694*
	(0.126)	(0.240)	(0.414)
Y = β_M			
ETF %	0.173***	0.0182	0.00879
	(0.0282)	(0.0185)	(0.00765)
Wtd SD	0.0224	0.712	2.168***
	(0.323)	(0.608)	(0.540)
Wtd turnover	0.253**	0.311**	0.735***
	(0.116)	(0.130)	(0.176)
Y = $sqrt(R^2)$			
ETF %	0.0660***	0.00353	0.00175
	(0.00832)	(0.00587)	(0.00155)
Wtd SD	0.187*	0.193	0.255
	(0.0941)	(0.164)	(0.173)
Wtd turnover	0.114***	0.0634*	0.155***
	(0.0291)	(0.0369)	(0.0418)

Table 6

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 %		
Mean β tercile 1	0.0272**	0.0061
	(0.013)	(0.009)
Mean β tercile 3	-0.0216*	-0.0029
	(0.013)	(0.016)
Difference	0.0488***	0.0089
	(0.010)	(0.008)
Terciles by Wtd SD		
Mean β tercile 1	0.0489***	-0.0008
	(0.008)	(0.007)
Mean β tercile 3	-0.0376**	-0.0038
	(0.016)	(0.014)
Difference	0.0864***	0.0030
	(0.012)	(0.011)
Terciles by Wtd turn		
Mean β tercile 1	0.0656***	0.0054
	(0.009)	(0.007)
Mean β tercile 3	-0.0579**	-0.0071
	(0.016)	(0.014)
Difference	0.1235***	0.0125
	(0.012)	(0.010)

Table 7

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. 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 %	-0.00532**	0.000582
	(0.00211)	(0.00258)
Wtd SD	-0.133	-0.216
	(0.347)	(0.165)
Wtd turn	-0.313***	-0.122***
	(0.0640)	(0.0425)
Stock turn	0.884***	-0.279
	(0.271)	(0.185)
Log(Mkt cap)	-0.0290	-0.0254
	(0.0198)	(0.0175)
B/M	0.0105*	0.00312
	(0.00614)	(0.00323)
S&P 500	-0.112***	0.0220
	(0.0336)	(0.0216)
Index %	0.0191***	0.00909**
	(0.00362)	(0.00384)
MF %	-0.000409	-0.000350
	(0.000471)	(0.000353)
Ins. %	-0.000480*	2.28e-05
	(0.000258)	(0.000217)
Time FE	YES	YES
Stock FE	YES	YES
N	186,860	186,860
R^2	0.0312	0.0257

Table 8

Stock sentiment beta. Month/stock panel regressions examine how a stock's exposure to ETF activities affects its sentiment beta. The sentiment beta is defined as the stock return exposure to the orthogonalized change in monthly Baker and Wurgler (2006) sentiment index, which is available through 2010.

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. 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)	(3)	(4)
		$Y = \beta_{SENT}$		
ETF %	0.00122			0.00105
	(0.00150)			(0.00146)
Wtd SD		0.188*		-0.0237
		(0.0717)		(0.0136)
Wtd turn			0.0548**	0.0580**
			(0.0138)	(0.0137)
Stock turnover	0.0367	0.0396	0.0410	0.0325
	(0.171)	(0.165)	(0.164)	(0.172)
Log$^{\text{(Mkt cap)}}$	-0.00138	-0.000630	-0.000628	-0.000856
	(0.00135)	(0.00110)	(0.00116)	(0.00140)
B/M	-0.00104	-0.000774	-0.000700	-0.000713
	(0.000833)	(0.000908)	(0.000928)	(0.000919)
S&P 500	-0.00734*	-0.00763*	-0.0102**	-0.00939**
	(0.00336)	(0.00305)	(0.00227)	(0.00244)
MF %	0.000124	0.000160**	0.000158**	0.000122
	(8.10e-05)	(5.73e-05)	(5.54e-05)	(8.78e-05)
Ins. %	-7.38e-05	-7.87e-05	-9.47e-05	-0.000105
	(6.65e-05)	(6.77e-05)	(6.94e-05)	(7.80e-05)
Year FE	YES	YES	YES	YES
Observations	14,718	14,718	14,718	14,718
R-squared	0.039	0.041	0.043	0.043

Table 9

Stock sentiment beta: Subsets by size and turnover. Coefficients from multivariate regressions on monthly data from July 2006 to December 2010. *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 control variables in Table 8 (*Stock turnover, Log(Market cap), B/M, S&P 500, MF% and Ins%*) 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) standard errors are used, and ***, ** and * signify statistical significance at the 1%, 5% and 10% levels.

Panel A

Y = β_{SENT}	Size 1	Size 2	Size 3
ETF %	0.00409*	8.23e-05	-0.000484
	(0.00183)	(0.00177)	(0.000650)
Wtd SD	0.0584	0.0454	0.0218
	(0.0844)	(0.0599)	(0.0570)
Wtd turnover	0.0447*	0.0225*	0.0333**
	(0.0197)	(0.00989)	(0.00842)

Panel B

Y = β_{SENT}	Turnover 1	Turnover 2	Turnover 3
ETF %	0.00421	0.00116	0.000206
	(0.00264)	(0.00229)	(0.00107)
Wtd SD	0.294**	0.0531	0.0651
	(0.0919)	(0.0565)	(0.0623)
Wtd turnover	0.0963***	0.0219	0.0331**
	(0.0185)	(0.0126)	(0.00729)

Table 10

Coefficients on weighted turnover for 6 months before and 6 months after the May 6, 2010 "flash crash". The "pre" period is 11/2009-4/2010. The "post" period is 6/2010-11/2010. *S Ratio* 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. ρ_M is the R-squared from the regression of daily excess stock returns on excess market 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. *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. 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.

	(1)	(2)	(3)	(4)	(5)	(6)
	$Y = Sratio$		$Y = \beta_M$		$Y = \rho_M$	
	Pre	Post	Pre	Post	Pre	Post
Wtd turn	1.198***	-1.046***	0.630**	-0.259	0.147*	-0.190**
	(0.0897)	(0.221)	(0.180)	(0.206)	(0.0622)	(0.0590)
Stock turn	-4.601***	-5.782**	-1.224**	-0.563	-1.570***	-1.351***
	(0.675)	(1.823)	(0.405)	(0.316)	(0.145)	(0.130)
Log(Mkt cap)	0.0536	-0.529	0.118	-0.100	0.00567	-0.0578
	(0.124)	(0.303)	(0.127)	(0.137)	(0.0355)	(0.0303)
B/M	0.0128	-0.246**	0.0103	-0.0759**	-0.00631	-0.0444***
	(0.0372)	(0.0698)	(0.0236)	(0.0253)	(0.0112)	(0.00951)
S&P 500	-0.454*	7.645***	-0.0826	0.762***	-0.0760	0.194***
	(0.182)	(0.338)	(0.139)	(0.120)	(0.0499)	(0.0275)
Index %	0.0256	-0.0431	0.00924	-0.0280*	0.00597*	-0.00470
	(0.0259)	(0.0392)	(0.00777)	(0.0126)	(0.00268)	(0.00733)
MF %	-0.00119	0.00226	1.22e-05	0.00768***	-0.00128**	0.00127*
	(0.00441)	(0.00139)	(0.00145)	(0.00133)	(0.000392)	(0.000620)
Ins. %	0.00170	-0.000552	0.00258**	0.00605***	0.000401	0.000625***
	(0.00260)	(0.00301)	(0.000995)	(0.00128)	(0.000207)	(0.000103)
Time FE	YES	YES	YES	YES	YES	YES
Stock FE	YES	YES	YES	YES	YES	YES
N	14,652	14,439	14,652	14,439	14,652	14,439
R^2	0.1221	0.1617	0.0659	0.0303	0.1142	0.1978