informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–26
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

ETF Activity and Informational Efficiency of Underlying Securities

Lawrence Glosten,[a] **Suresh Nallareddy,**[b] **Yuan Zou**[c]

[a] Columbia University, New York, New York 10027; [b] Duke University, Durham, North Carolina 27708; [c] Harvard Business School, Boston, Massachusetts 02163

Contact: lrg2@gsb.columbia.edu (LG); suresh.nallareddy@duke.edu, https://orcid.org/0000-0003-0957-322X (SN); yzou@hbs.edu (YZ)

Received: April 3, 2017
Revised: June 26, 2018
Accepted: June 10, 2019
Published Online in Articles in Advance: April 27, 2020

https://doi.org/10.1287/mnsc.2019.3427

Abstract. This paper investigates the effect of exchange-traded funds' (ETFs') activity on the short-run informational efficiency of their underlying securities. We find that ETF activity increases short-run informational efficiency for stocks with weak information environments. The increase in informational efficiency results from the timely incorporation of systematic earnings information. In contrast, we find no such effect for stocks with stronger information environments. ETF activity increases return comovement, and this increase is partly attributable to the timely incorporation of systematic earnings information. Further, ETF activity is associated with an attenuation of postearnings-announcement drift and an increase in active share lending.

History: Accepted by Suraj Srinivasan, accounting.

Keywords: ETF • informational efficiency • co-movement • systematic earnings • post-earnings-announcement drift • share lending

1. Introduction

The asset-management industry has witnessed tremendous growth in exchange-traded funds (ETFs). As a result, roughly 30% of U.S. equity trading volume is attributable to ETFs (Boroujerdi and Fogertey 2015).[1] The literature on the consequences of ETFs is in its early stages (Lettau and Madhavan 2018), and the early evidence suggests that ETFs may have distorted the capital markets (e.g., Wurgler 2010, Ramaswamy 2011, Da and Shive 2016, Antoniou et al. 2018, and Ben-David et al. 2018). Further, the Securities and Exchange Commission (SEC) has called for more research and discussion on the consequences of ETFs.[2] In this paper, we investigate the effect of ETF activity on the short-run informational efficiency of their underlying securities. We find that ETF activity increases the informational efficiency by improving the link between short-run fundamentals and stock prices. Specifically, firms with more ETF activity reflect incrementally more earnings information in their current stock returns.

In the absence of ETFs, as information arrives, investors must assess the implications of that information for each security. As a result, information might not be reflected in some segments of the market (e.g., firms with weak information environments, low liquidity, or significant short-sale constraints) on a timely basis. However, in the presence of ETFs, because investors have the ability to trade a basket of securities as opposed to individual stocks, information could be reflected in a more timely manner for a broader cross-section of stocks. This feature could result in an improved link between fundamentals and stock prices, particularly for stocks that are difficult to trade. For example, during the recent financial crisis (September and October of 2008), the SEC banned short selling for 797 financial stocks. However, all ETFs were explicitly exempted from the ban (Karmaziene and Sokolovski 2014). Further, short sellers can employ ETFs to circumvent short-sale constraints (e.g., Li and Zhu 2016).[3] Furthermore, stocks that experience an increase in passive index ownership are associated with greater public scrutiny and enhanced corporate governance (e.g., Boone and White 2015 and Appel et al. 2016). These factors could increase the informational efficiency of stocks as the passive ownership increases. Finally, we discuss theoretical support for ETF activity leading to information being reflected in underlying securities.

On the other hand, ETF activity could transmit potential nonfundamental shocks (e.g., sentiment-related mispricing), resulting in a breakdown of the link between fundamentals and stock returns (Da and Shive 2016, Ben-David et al. 2018). Also, ETF ownership on average is less than 5% of the shares outstanding. Such low ownership may not affect the informational efficiency of underlying securities. Therefore, the effect of ETF activity on informational efficiency is an open question.

ETF activity for a stock is defined as the quarterly change in ETF ownership. ETF activity could be a result of excess demand and supply from the investors or ETF additions/deletions (Abner 2010). Using a large cross-section of ETF holdings data from January 2004 to December 2013, we document that an increase in ETF activity is accompanied by an increase in the short-run informational efficiency of the underlying stocks, as reflected in the increase in the relation between stock returns and earnings news. The effect of ETF activity on short-run informational efficiency should be conditional on the information environment. Consistent with expectations, when we conduct the informational-efficiency tests within different segments of the market, we find significant and improved short-run informational efficiency among small firms (firms with market capitalization below the NYSE 50th percentile) and stocks with low analyst following (firms with analyst following below the 75th percentile).[4] In contrast, we are unable to document such improvement for big firms and stocks with high analyst following.

Next, we expect systematic information that affects a basket of securities to result in ETF activity. In particular, as our model shows, an investor with systematic information is most likely to use ETFs to speculate because that will incur lower transaction costs. This speculation is likely to generate ETF activity and help stock prices reflect systematic information. On the other hand, a speculator with firm-specific information might buy or sell the firm's stock, using ETFs to hedge the systematic risk. However, speculators doing this are unlikely to generate ETF activity because across the stocks in the ETF, roughly half will generate positive information leading to ETF sells, and half will generate negative information leading to ETF purchases. Therefore, if ETF activity results in increased informational efficiency for underlying stocks, then the increase in informational efficiency should be attributable to systematic information rather than idiosyncratic information. This mechanism is different compared with that in the literature that investigates the effect of institutional investors on the informational efficiency (e.g., Jiambalvo et al. 2002 and Piotroski and Roulstone 2004). In particular, institutional investors' investment decisions are influenced by private firm-specific information, and, hence, changes in their holdings reflect private information (Chakravarty 2001). We find evidence consistent with the conjecture that ETF activity results in prices that reflect systematic information in a timely manner rather than idiosyncratic earnings information, resulting in increased short-run informational efficiency. In particular, we decompose earnings into their systematic and firm-specific components and find that the commonality component of earnings explains the increase in short-run informational efficiency but not the idiosyncratic firm-level earnings.

The literature finds that index membership increases comovement, and this increase is driven by nonfundamental factors (Harris and Gurel 1986, Vijh 1994, Barberis et al. 2005, Peng and Xiong 2006, Da and Shive 2016). However, by making it easier to trade stocks with similar characteristics, ETF activity could potentially move prices to reflect more systematic information and could contribute to higher comovement. Therefore, increases in comovement could also be driven by fundamental information. Consistent with expectations, we find that the increase in return comovement is partially explained by systematic earnings information.

Further, we investigate the relation between ETF activity and postearnings-announcement drift (PEAD). If ETF activity incorporates information in a timely manner into underlying stock prices, then PEAD should be attenuated. Consistent with expectations, we find that ETF activity is associated with lower PEAD strategy returns. Specifically, the difference in five-factor adjusted returns to the PEAD strategy between low and high ETF activity is 0.88%, and it is statistically significant at the 1% level. This evidence suggests that ETF activity incorporates information quickly into stock prices, which results in lower PEAD strategy returns as the ETF activity increases.

Next, employing Russell 1000/2000 index reconstitution as a setting, we investigate the effects of ETF activity on the short-run informational efficiency of underlying securities, and we corroborate our main findings. Russell 1000/2000 index reconstitution offers a setting in which firms with similar characteristics have a significant variation in ETF ownership. A market capitalization rank of 1,000 is the cutoff rank for the Russell 1000 index; the next 2,000 ranked stocks constitute the Russell 2000 index. Chang et al. (2015) document that a small market capitalization change around the market capitalization rank of 1,000 moves a stock between the Russell 1000 and 2000 indices, but the firms close to this rank remain similar in terms of firm characteristics. As the indexes are value-weighted, if one of the lowest-market-capitalization stocks of the Russell 1000 index switches to the Russell 2000 index, ETF ownership of that stock increases significantly. Using this setting and a difference-in-differences design, we find that the change in short-run informational efficiency for firms moving from the Russell 1000 to the Russell 2000 is positive and significantly greater than the change in short-run informational efficiency for firms moving from the Russell 2000 to the Russell 1000 index. Further, increases in short-run informational efficiency are significantly greater for the firms with weak information environments when these firms are switched to the Russell 2000 index than when they are switched to the Russell 1000 index. It is worth noting that the evidence using this setting is only suggestive, as concurrent changes in other firm policies and the institutional environment could potentially affect the short-run informational efficiency.

Like our paper, the work of Israeli et al. (2017) investigates the implications of ETF activity for informational efficiency. However, our paper differs from theirs in several ways. First, Israeli et al. (2017) document that ETF activity leads to the deterioration of pricing efficiency in the long run. In contrast, we investigate whether ETF activity incorporates current-quarter earnings news into current-quarter stock returns in a timely manner. Our premise is that ETF activity does not predict future fundamentals, but merely incorporates earnings news (e.g., sector-level news and macroeconomic news) in a timely manner. We show that ETF activity does not, in fact, predict future earnings and stock returns. Second, we investigate the channels through which short-run informational efficiency is improved by ETF activity. In particular, we investigate the role of both alternative information channels and alternative ETF activity channels. Third, we investigate the effect of ETF activity on short-run informational efficiency for different segments of the market. Fourth, we investigate the relation between ETF activity and postearnings-announcement drift. Finally, we investigate the link between ETF activity and active share lending. We find that ETF activity is positively related to active share lending, which may reduce shorting costs.

This paper makes several contributions to the literature. First, we describe a model of factor-informed investors and idiosyncratically informed investors choosing to trade in the ETF and the underlying securities. The model suggests that factor-informed investors will frequently trade both the ETF and the underlying securities. In contrast, idiosyncratically informed investors are likely to trade only the security about which they have information. This motivates our examination of the extent to which ETF activity in a stock helps reveal systematic as opposed to idiosyncratic information.

Second, our paper contributes to the growing debate on the consequences of index-linked products for the stock market (see Madhavan 2016 and Lettau and Madhavan 2018 for a review). Specifically, we document whether, how, and when ETF activity increases the short-run informational efficiency of underlying stocks. ETF activity increases the short-run informational efficiency for firms with weak information environments by incorporating systematic accounting information into stock prices in a timely manner. In contrast, we find no such effect for firms with stronger information environments. Third, prior literature documents that index membership increases comovement (Barberis et al. 2005, Greenwood and Sosner 2007, Staer 2014, Da and Shive 2016). Our findings document that this increase is partly attributable to the timely incorporation of fundamental information into stock prices and not fully driven by nonfundamental factors.

Fourth, we document that ETF activity is associated with attenuation of returns to the PEAD trading strategy. Fifth, we document that ETF activity is positively related to active share lending, which may reduce stock shorting costs. Sixth, by documenting the effect of ETF activity on the informational efficiency of the underlying stocks, we provide evidence in support of the long-standing prediction that policies that stimulate liquidity and ameliorate trading costs improve market efficiency (Chordia et al. 2014). Finally, financial regulators are concerned about the impact of ETF activity on liquidity, volatility, and informational efficiency. In response to regulators' concerns about the potential consequences of ETFs for the capital markets, we provide evidence that ETF activity, in fact, increases short-run informational efficiency for some segments of the equity markets.

The rest of the paper is organized as follows. Section 2 provides institutional details and theoretical predictions. Section 3 describes the data and the main variable construction. Section 4 presents the empirical results. Section 5 offers concluding remarks.

2. Institutional Details, Theoretical Predictions, and Related Literature

2.1. Institutional Details

ETFs are a hybrid of two antecedents, mutual funds and investment trusts. Like mutual funds, ETFs are open-ended funds that can create and redeem shares at any time. Like investment trusts, but unlike mutual funds, ETFs are traded on organized stock exchanges throughout the day, whereas open-ended mutual funds can be bought or sold only at the end of the day for net asset value (NAV). ETFs provide investors access to diversified portfolios in a less expensive and more convenient way than traditional mutual funds do. For example, the average expense ratio is around 0.25% per year for ETFs, whereas it is around 1% for mutual funds (Nallareddy and Ogneva 2017).[5] Please refer to Lettau and Madhavan (2018) for an excellent introduction to and review of ETFs.

The unique creation/redemption mechanism associated with ETFs ensures that ETF shares will expand or contract based on demand from investors. In the primary market, only authorized participants (APs), who are large broker-dealers, buy from and sell to the ETF sponsor large blocks of ETF shares. In the secondary market, investors can then buy and sell ETF shares just like common stocks. Because the price of ETF shares is determined by the demand and supply on the secondary market, the price is not always equal to the NAV. APs try to ensure that intraday prices approximate the NAV of the underlying assets through the creation/redemption of ETF shares. For example, if there is an increased demand for the ETF shares, the APs can buy a block of new shares of the ETF, called "creation units," from the ETF sponsor by transferring the basket of the securities to the sponsor, and then they can sell the new ETF shares on the secondary market. Importantly, the creation/redemption mechanism of ETFs on the primary market indicates excess demand for ETF shares from investors. In other words, an increase in the shares of an ETF reflects an increase in the size of the ETF portfolio. This implication helps us build our proxy for ETF activity.

In addition, ETFs provide investors access to stocks that were previously hard to trade. For example, a small-cap ETF, VB, is based on small-capitalization stocks, for which the underlying stocks are less liquid. However, VB holds close to $21.6 billion in fund net assets (as of December 31, 2017) and trades at very low cost. Such benefits of trading ETFs attract traders in the secondary market.

2.2. Theoretical Predictions

There have not been many modeling efforts involving the trading of an ETF and the underlying stock while dealing with the option to invest in firm-specific information, factor information, or no information. Existing efforts have been recent. Glasserman and Mamaysky (2016) use a rational expectations model to look at the effect of ETF trading. Interestingly, the model concludes that, in equilibrium, those who are informed about the underlying factor take the counter position to the net noise trade. This is required by the law of one price: In equilibrium, the price of the ETF must be equal to the weighted sum of the prices of the underlying security. Consequently, the informed traders may have positive or negative news, but the trades they make in equilibrium are orthogonal to the news they received. This also means that, in equilibrium, the number of ETF shares outstanding will not change—the market for ETF shares clears even though it does not have to.

This result suggests that investigating a Kyle (1985) type model might be fruitful, and this is done by Bhattacharya and O'Hara (2018) as well as Cong and Xu (2016). The first paper assumes that prices in the ETF and underlying securities are a function only of the order flow in each individual security. However, after the first round of trade, informed traders can undo their position (or not) at a price that does reflect all the information in the market, but with no price impact. As a consequence, there is sometimes a herding equilibrium in which the informed get out of their position at these intermediate prices. This, the paper shows, leads to enhanced transmission of noise trade and divergence of the ETF price from the portfolio cost.

If we were to adopt a different construct and assume that prices in the ETF and the underlying securities are determined by all observable order flows, then the prices of the ETF and the underlying securities do satisfy the law of one price because all available information is used to price all securities. However, this model proves to be unsuitable. When traders informed about the fundamental factor mimicked by the ETF are allowed to trade in both the underlying security and the ETF, equilibrium will require a very specific relation between the order flow sensitivities of the individual securities and the ETF order flow sensitivity of the underlying security in order to avoid unbounded expected profits. Furthermore, the required relation between the order flow intensities is likely inconsistent with zero profit to liquidity suppliers. That is, equilibrium fails to exist.[6]

We model a small part of the financial market by looking at an ETF targeting a factor F and do so by creating a portfolio of N underlying securities with payoffs $F + \varepsilon_i$, where the idiosyncratic terms are independent of the factor F. The factor and each idiosyncratic term take values in $\{0,1\}$. Each security has one share outstanding, so the fund acquires a fractional share $1/N$ of each. The payoff on the ETF, per share, is thus

$$F + \frac{1}{N}\sum_i \varepsilon_i. \tag{1}$$

We suppose that there are individuals informed about the factor, F, and other individuals informed about the idiosyncratic terms. Without loss of generality, we take the individuals' information to be perfect. Traders arrive one at a time at the markets for the individual securities and the market for the ETF. That is, we consider Glosten/Milgrom-type markets for the securities (Glosten and Milgrom 1985). There are, of course, orthogonal noise traders in all of the markets who buy or sell with equal probability.

There are $N + 1$ markets, corresponding to the N underlying securities, labeled $1,\ldots,N$, and the ETF market labeled E. There are $2N + 2$ information types—$N + 1$ uninformed noise types (labeled Z), N idiosyncratically informed types, and 1 factor-informed type. We denote by α_{ij} the probability that the next arrival for trade in market j, $j = E,1,\ldots,N$, has information i, $i = F, 1,\ldots,N, Z1,\ldots,ZN, ZF$. At a point in time, just before a transaction in any one market, let the expected values of F and ε_i be P_F and P_i, respectively. It is obvious that an agent informed about ε_i cannot trade profitably in market j, so to simplify the notation, we use α_i rather than α_{ii}. It is possible but unlikely that an i-informed trader might trade in the ETF. Then, standard calculations yield the bid and ask in the ETF market:

$$B_E = P_F + \frac{1}{N}\sum P_i - \frac{\alpha_{FE}P_F(1-P_F) + \frac{1}{N}\sum \alpha_{iE}P_i(1-P_i)}{\alpha_{FE}(1-P_{F)} + \sum \alpha_{iE}(1-P_i) + .5\alpha_{ZE}}, \quad (2)$$

$$A_E = P_F + \frac{1}{N}\sum P_i + \frac{\alpha_{FE}P_F(1-P_F) + \frac{1}{N}\sum \alpha_{iE}P_i(1-P_i)}{\alpha_{FE}P_F + \sum \alpha_{iE}P_i + .5\alpha_{ZE}}. \quad (3)$$

Similarly, the bid and ask in market i are given by:

$$B_i = P_F + P_i - \frac{\alpha_{Fi}P_F(1-P_F) + \alpha_i P_i(1-P_i)}{\alpha_{Fi}(1-P_F) + \alpha_i(1-P_i) + .5\alpha_{zi}}, \quad (4)$$

$$A_i = P_F + P_i + \frac{\alpha_{Fi}P_F(1-P_F) + \alpha_i P_i(1-P_i)}{\alpha_{Fi}P_F + \alpha_i P_i + .5\alpha_{zi}}. \quad (5)$$

It is an equilibrium for any i-informed agent that has good news to refrain from buying in the ETF market if

$$\frac{1}{N}(1-P_i) - \frac{\alpha_{FE}P_F(1-P_F)}{\alpha_{FE}P_F + .5\alpha_{ZE}} \le 0. \quad (6)$$

Although it is certainly possible that the above expression could be positive, it is unlikely. Conditional on an i-informed agent having good news, it is likely that other i-informed agents have already traded on their information, increasing P_i and making the first term small. The following calculations are reasonable and conservative: with $\alpha_{ZE} = 0.9$ and $P_F = P_i = 0.5$, it is immediate that i-informed traders will not trade in the ETF if the ETF has 10 names or more. We are confident that Equation (6) will hold, and trade in the ETF will be dominated by noise and factor-informed traders. Thus, trading in the ETF should not reveal idiosyncratic information.[7]

On the other hand, it is quite likely that factor-informed investors will trade in the underlying securities as well as in the ETF, particularly soon after the factor information has been discovered. The condition that must be satisfied for a factor-informed buyer to buy in market i is

$$.5\alpha_{Zi}(1-P_f) + \alpha_i(P_i - P_F) > 0. \quad (7)$$

When values P_i and P_F are close to 0.5, this is clearly satisfied. Conditional on positive news about the factor, however, P_F will increase, it will be less likely to be the case, and factor-informed trading will occur only in the ETF market.

With these observations, the spread in the ETF and the underlying will be when all values are 0.5

$$S_E = \frac{\alpha_{FE}}{\alpha_{FE} + \alpha_{ZE}}, \quad (8)$$

$$S_i = \frac{\alpha_{Fi} + \alpha_i}{\alpha_{Fi} + \alpha_i + \alpha_{zi}}. \quad (9)$$

Thus, the underlying asset will face adverse selection from both factor-informed and idiosyncratically informed agents. We conclude that the spread will be wider in the underlying than in the ETF, even if the extent of noise trading is the same in the underlying and the ETF.

The model allows some speculation regarding how ETF introduction will affect the spread in the constituent parts. It is likely that noise trade will decrease in the underlying, going to the ETF. This is because partial portfolio liquidation (or increased saving) is easier using the ETF. Thus, the denominator in (9) will decrease. Furthermore, the lower spread in (8) may attract more searches for factor information, increasing the numerator in (9). On the other hand, the increase in the spread in (9) will make the search for idiosyncratic information less profitable. To say more would require more data on the costs of information acquisition.

The analysis does suggest that if ETF ownership of a firm's shares increases, that would be associated with factor-informed traders buying the ETF. There should be a positive association between that ownership and the stock return in the period in which the ownership increases. We do, in fact, discover this in the data, though the effect is limited to smaller firms with less analyst following.

There is, of course, another reason for changes in the percentage of a firm's shares owned by ETFs: ETF rebalancing, particularly adding and subtracting portfolio firms. A mental experiment will illustrate why a firm's addition to an ETF might improve that firm's information environment. Even if the ETF to which the firm is being added is closely related to other ETFs that the firm is in, this simple addition increases the number of quotes that a factor-informed investor can check by 12. It is quite likely, given random idiosyncratic fluctuations in liquidity, that an investor can find better terms of trade. This has two effects—given that an investor has information, that information is more likely to be traded on profitably. However, this also increases the potential profitability of future searches for information, thereby increasing the amount of information gathering.

There is a second reason that being added to an ETF could change the informational efficiency of the market for a firm's shares, particularly if the firm is small. Reference to a standard ETF prospectus shows that the typical ETF will engage in share lending to short sellers. Thus, if a firm is added to one ETF without being dropped from another, the number of shares available for shorting goes up, possibly reducing the cost of shorting those shares. This, as above, is useful for speculators because it reduces the cost of a negative bet. Because investors looking for information will encounter bad news roughly half of the time, the increase in the return to investing in information will increase information gathering. We verify this assertion in Section 4.9 and document a positive association between ETF activity and share lending.

To conclude, we have some theoretical support for ETF creation/redemption leading to information being reflected in a firm's stock price. We have reason to believe (though we lack a rigorous model) that a firm's addition to an ETF enhances that firm's information environment, whereas being deleted has a deleterious effect. This analysis is contrary to the more behaviorally oriented prediction that ETFs harm informational efficiency.

2.3. Related Literature

A number of studies document the negative effects of ETFs. Ramaswamy (2011) links the rise of ETFs to greater systemic risk. Hamm (2014) documents that ETF ownership is positively related to a stock's illiquidity. Ben-David et al. (2018) and Krause et al. (2014) provide evidence that the arbitrage activity between ETFs and the underlying stocks leads to an increase in intraday and daily stock volatility because of the transmission of liquidity shock from ETFs to the underlying stocks. Da and Shive (2016) document that higher ETF arbitrage activity contributes to return comovement at both the fund and the stock levels. Bhattacharya and O'Hara (2018) show that ETFs increase instability and herding. The main conclusion that we can draw from this literature is that nonfundamental demand shocks might be transferred from the ETFs to the underlying securities.

A number of studies also highlight the positive effects of ETF activity. Hasbrouck (2003) documents that ETFs improve intraday price discovery for the underlying stocks during the sample period March 2000 to May 2000. Boehmer and Boehmer (2003) document that the initiation of three ETFs on the NYSE increased liquidity and market quality. In contrast to the intraday studies, our paper covers a much broader cross-section of ETFs and stocks. It also has a longer time period, which allows us to examine the broader consequences of ETF activity, particularly given the increasing popularity of ETFs since 2000. Further, unlike previous studies, we investigate whether ETF activity incorporates fundamental information into stock prices in a timely manner.

Our paper is related to the literature on institutional ownership (IO) and price discovery (e.g., Jiambalvo et al. 2002 and Piotroski and Roulstone 2004), but the findings using the institutional ownership setting do not extend to the ETF setting for several reasons. First, institutional investors do fundamental analysis and pick stocks with specific characteristics, such as firms with better disclosures and information environments (e.g., Diamond and Verrecchia 1991, Healy et al. 1999, Bushee and Noe 2000, Gompers and Metrick 2001, Lambert et al. 2007, and Ferreira and Matos 2008). In contrast, ETF investors do not have such flexibility. Second, institutional investor investment decisions are influenced by private firm-specific information, so changes in their holdings convey private information (Chakravarty 2001). ETF activity, however, is related to common information, as ETFs include a basket of securities. Therefore, unlike institutional investor activity, ETF activity likely reflects market and industry information rather than firm-level idiosyncratic information. Third, given the sophistication and sizable stakes of institutional investors, they have incentives to collect and act on such information, which results in informational efficiency (Shleifer and Vishny 1986, 1997; Gillan and Starks 2003). Finally, ETF activity and IO activity have little correlation in the cross-section. Specifically, the Pearson (Spearman) correlation between ETF activity and IO activity is 0.006 (0.008), and only the Spearman correlation coefficient is statistically significant. Nevertheless, in all our analyses, we control for IO activity when investigating the relation between ETF activity and informational efficiency.

It is also important to note that the arguments and findings in our paper apply to index mutual funds and some futures contracts (e.g., Standard & Poor's (S&P) mini futures and industry futures). However, institutions are increasingly using ETFs in place of equity futures contracts (Greenwich Associates 2016). Specifically, 40% of the institutions in year 2015 used ETFs in place of equity futures contracts, and 78% of the futures users planned to replace them with ETFs within the next 12 months (Greenwich Associates 2016). Further, ETFs are a better investment alternative for fully funded investors, because ETFs have lower transaction costs and do not face the mispricing risk often associated with futures' roll dates (Madhavan et al. 2014). It is true that factor-informed traders can and probably do use index mutual funds. Still, as many have argued, ETFs are even easier to use. In particular, the ability to exit a position at any time, rather than having to wait until the end of the day, makes risk management of a bet much easier.

3. Data and Variable Construction

3.1. Data

We obtain ETF data from the Center for Research in Security Prices (CRSP) daily stock file, using the share code of 73, which uniquely identifies ETFs in the CRSP universe.[8] Quarterly ETF holdings data are from the Thomson–Reuters Mutual Fund holding database (S12). We merge the holding data with the ETF data using the MFLINKS tables. This procedure yields the final sample of 447 ETFs, where each ETF has the holdings data for each stock for the quarters from

2004 to 2013.[9] Stock return and accounting data are from the intersection of the CRSP and Compustat datasets from 2004 to 2013. Our sample includes firms listed on the NYSE, AMEX, or NASDAQ that have CRSP share codes 10 or 11. To align ETF holding data with firm-level accounting data, we include only firms with fiscal-year ends in March, June, September, or December. Further, we exclude stocks with prices less than \$2 to mitigate market microstructure noise. In addition, institutional ownership could be related to informational efficiency (Boehmer and Kelley 2009); therefore, we control for institutional ownership in all of our analyses. Therefore, our final sample consists of firms with institutional ownership data available from the Thomson–Reuters Institutional Holdings (13F) Database. Finally, to alleviate the effects of outliers, we winsorize all independent variables at the 1% and 99% levels. The final sample contains 78,984 firm-quarters.

3.2. ETF Activity

We use changes in a stock's ETF ownership as a proxy for ETF activity in each stock, because it is a direct measure that aggregates the net activity of a stock from different ETFs. ETF ownership is calculated as the proportion of shares owned by all the ETFs in the stock's total shares outstanding. Specifically, the ETF ownership for each stock and quarter, $ETF_{i,t}$, is calculated as

$$ETF_{i,t} = \frac{\sum_{j=1}^{J} SHARES_{j,t}}{Total\ Shares\ Outstanding_{i,t}}, \tag{10}$$

where j is the set of ETFs holding stock i; $SHARES_{j,t}$ is the number of stock i's shares held by ETF j at the end of quarter t; and $TotalSharesOutstanding_{i,t}$ is the total shares outstanding for stock i at the end of quarter t. All the variables are measured at the end of each quarter. The change in ETF ownership is calculated as the quarterly difference in $ETF_{i,t}$. To mitigate concerns of outliers and to interpret coefficient estimates, we convert the change in $ETF_{i,t}$ into a rank variable. Specifically, for each quarter, we sort stocks based on change in $ETF_{i,t}$ and rank this variable into 10 groups [1, 10]. We then divide the rank variables by 10, such that the ETF activity covering stock i, $ETF_{i,t}$, is between 0.1 and 1. Thus, $\Delta ETF_{i,t} = 0.1$ indicates the greatest decrease in ETF ownership in magnitude, whereas $\Delta ETF_{i,t} = 1$ indicates the greatest increase in ETF ownership.

Table 1, panels A and B, presents descriptive statistics. Both the number of ETFs and ETF ownership per stock have increased over time. The average stock in our sample is held by 19.32 ETFs as of the last quarter of 2013. The ETF ownership per stock increased from 1.2% in 2004% to 5% in 2013. Mean (median) ETF ownership is 3.6% (2.9%).

3.3. Earnings Information

We use seasonally adjusted earnings deflated by beginning-of-quarter price as a measure for the fundamental information. Specifically, seasonally adjusted earnings in quarter t are measured by

$$EARN_{i,t} = \frac{(X_{it} - X_{it-4})}{P_{it-1}}, \tag{11}$$

where X_{it} is earnings per share excluding extraordinary items for firm i in quarter t, and P_{it-1} is price per share for firm i at the end of quarter $t-1$. Figure 1 presents the timeline of our variable measurement. Earnings information for quarter t is released in quarter $t+1$, and it is unavailable at the end of quarter t.

4. Empirical Analyses

4.1. ETF Activity and the Contemporaneous Returns–Earnings Relation

In this section, we examine whether ETF activity affects the extent to which fundamental earnings information is incorporated into underlying stock prices. To do so, we estimate the following Fama and MacBeth (1973) regression:

$$\begin{aligned} Ret_{i,t} = {} & b_{0,t} + b_{1,t}Earn_{i,t} + b_{2,t}\Delta ETF_{i,t} + b_{3,t}Earn_{i,t} \times \Delta ETF_{i,t} \\ & + b_{4,t}Inst\_residual_{i,t} + b_{5,t}Earn_{i,t} \times Inst\_residual_{i,t} \\ & + b_{6,t}MTB_{i,t-1} + b_{7,t}Size_{i,t-1} + b_{8,t}STD_{i,t-1} \\ & + b_{9,t}Ret_{i,(t-12,t-2)} + b_{10,t}Loss_{i,t} + b_{11,t}Earn_{i,t} \times Loss_{i,t} \\ & + b_{12,t}Earn_{i,t-1} + b_{13,t}ETF_{i,t-1} \\ & + b_{14,t}Earn_{i,t-1} \times \Delta ETF_{i,t} + b_{15,t}\beta_{i,t-1} \\ & + b_{16,t}Ret_{i,t+1} + \varepsilon_{i,t}, \end{aligned} \tag{12}$$

where $Ret_{i,t}$ is the stock return for stock i during quarter t, and $\Delta ETF_{i,t}$ is the within-quarter rank for quarterly change in ETF ownership, scaled to [0.1,1]. $Earn_{i,t}$ is seasonally adjusted quarter earnings. Earnings information is realized after the quarter-end and hence unavailable in real time. The objective of using unavailable earnings information is to investigate whether ETF activity incorporates information about a current quarter's earnings that can be inferred from alternative information sources before management announces the earnings. The coefficient b_1 measures the relation between current returns and current earnings, and b_3 captures the effect of ETF activity on short-run informational efficiency. A positive b_3 would indicate that ETF activity pushes prices to reflect more fundamental information. To the extent that stock returns contain past earnings information, we also include $Earn_{i,t-1}$. We control for other characteristics that are either related to stock returns or shown to affect the returns–earnings relation: namely, $MTB_{i,t-1}$ (market-to-book ratio), the market value of equity to the book value of equity; $Size_{i,t-1}$, the natural

Table 1. Descriptive Statistics

Panel A: ETF ownership over time

Year	$ETF_{i,t}$	$\#ETF_{i,t}$
2004	1.222%	8.788
2005	1.610%	9.469
2006	1.916%	8.753
2007	2.215%	13.489
2008	2.720%	17.483
2009	3.000%	15.838
2010	2.487%	11.955
2011	4.323%	19.415
2012	4.542%	17.693
2013	4.925%	19.316

Panel B: Descriptive statistics of key variables of interest

Variable	N	Mean	Median	Q1	Q3	Standard deviation
$\Delta ETF_{i,t}$	78,984	0.002	0.001	−0.003	0.006	0.015
$ETF_{i,t}$	78,984	0.036	0.029	0.014	0.051	0.028
$Size_{i,t}$	78,984	6.970	6.817	5.692	8.096	1.726
$MTB_{i,t}$	78,984	2.793	1.975	1.261	3.274	3.902
$Loss_{i,t}$	78,984	0.227	0.000	0.000	0.000	0.419
$STD_{i,t}$	78,984	0.001	0.000	0.000	0.001	0.003
$Beta_{i,t}$	78,984	1.152	1.108	0.751	1.511	0.630
$Ret_{i,t}$	78,984	0.032	0.025	−0.088	0.137	0.227
$Earn_{i,t}$	78,984	0.000	0.001	−0.006	0.007	0.070
$Inst\%$	78,984	0.648	0.704	0.490	0.845	0.243
$\Delta Inst\%$	78,984	−0.014	0.001	−0.021	0.022	3.837
$\#Analyst_{i,t}$	78,984	7.629	5.667	2.333	11.000	6.907

Notes. Table 1 presents the descriptive statistics of ETF ownership as well as all other variables of interest. The sample spans 40 quarters from Q1:2004 to Q4:2013. Panel A presents ETF ownership over time. $ETF_{i,t}$ is the percentage of stock held by ETFs, computed for each year as an average across stocks and quarters. $\#ETF_{i,t}$ is the average number of ETFs that hold the stock. Panel B presents descriptive statistics for other key variables of interest. Variable definitions are in the appendix.

logarithm of the market value of equity at the beginning of the quarter; $STD_{i,t-1}$, the standard deviation of earnings during the past 20 quarters (5 years) preceding quarter t; $Ret_{i,(t-12,t-2)}$, stock returns compounded during 12 and 2 months in the preceding quarter t; $LOSS_{i,t}$, an indicator variable that equals 1 if quarterly earnings for firm i are negative, and 0 otherwise; the interaction of $Earn_{i,t}$ and $LOSS_{i,t}$; and $\beta_{i,t-1}$, the beginning-quarter market beta. We also control for the effect of the level of beginning-quarter ETF ownership, $ETF_{i,t-1}$, in our regressions, as the literature documents that it affects the volatility and bid–ask spreads of the underlying stocks (Da and Shive 2016, Israeli et al. 2017, Ben-David et al. 2018), which might, in turn, affect stock returns. Additionally, we control for the interaction of ETF activity and past earnings, $Earn_{i,t-1} \times \Delta ETF_{i,t}$, to control for the effect of ETF activity on the relation of returns and earnings during the last quarter. We also follow prior research to include future return, $Ret_{i,t+1}$, to account for the measurement error in using earnings in the earnings response coefficient regression (Collins et al. 1994, Lundholm and Myers 2002).

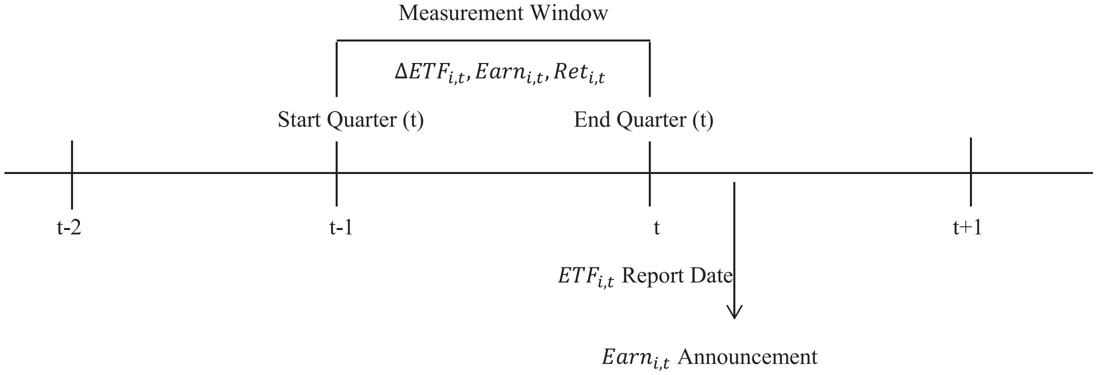
To control for the effect of institutional ownership on short-run informational efficiency, we include the quarterly change in institutional ownership in our analyses. However, because ETF activity could be correlated with institutional ownership, to isolate the effect of ETF activity, we orthogonalize change in institutional ownership with respect to change in ETF ownership. To do so, we adopt quarterly cross-sectional regressions as follows:

$$\Delta Inst_{i,t} = b_{0,t} + b_{1,t}\Delta ETF_{i,t} + Resid_{i,t}, \qquad (13)$$

where $\Delta Inst_{i,t}$ is the quarterly change in institutional ownership for stock i during quarter t; $\Delta ETF_{i,t}$ is the quarterly change in ETF ownership for stock i during quarter t; and $Inst\_residual_{i,t}$ is the residual from the above regression, capturing the change in institutional ownership that is not explained by ETF activity.[10] By doing so, we can isolate the effect of ETF activity on short-run informational efficiency that is not confounded by institutional ownership.

Table 2 presents the time-series average coefficients of the cross-sectional regression of returns on contemporaneous earnings. The evidence from Table 2

Figure 1. Timeline for Variable Measurement



suggests that ETF activity incorporates contemporaneous earnings information into stock prices, thereby increasing short-run informational efficiency. In contrast, institutional ownership that is orthogonalized to ETF activity does not increase short-run informational efficiency related to accounting information. Specifically, as shown in columns (1) to (7) of Table 2, the interaction between ETF activity and earnings information is statistically significant in explaining stock returns across different specifications, whereas the interaction between orthogonalized institutional ownership and earnings is not significantly related to stock returns.[11] Next, we find that ETF activity and stock returns are contemporaneously related. Also, consistent with expectations and prior literature, past and contemporaneous earnings are related to quarter t stock returns. The coefficients on control variables are consistent with expectation and prior literature. The negative coefficients on $Size_{i,t-1}$ and $MTB_{i,t-1}$ are consistent with the literature on the size effect (Banz 1981) and the growth effect (Chan et al. 1991, Lakonishok 1994). Negative coefficients on loss and the interaction between loss and earnings are also consistent with prior literature (Hayn 1995, Basu 1997). The insignificant coefficient on $ETF_{i,t-1}$ is consistent with Israeli et al. (2017). The insignificant coefficient estimate on $Ret_{i,(t-12,t-2)}$ is consistent with findings that earnings momentum subsumes price momentum (Chordia and Shivakumar 2006).

Overall, Table 2 provides evidence that ETF activity increases the returns–earnings relation, suggesting that short-run informational efficiency is improved. We conjecture that because ETFs enable investors to trade a basket of securities, ETF activity reflects accounting information into a broader cross-section of stocks. In contrast, in the absence of ETF activity and as information arrives, investors have to assess the implications of information for each security. As a result, information might not be reflected in some segments of the market (e.g., firms with weak information environments, low liquidity, or short-sale constraints) on a timely basis.

To provide further evidence in support of our conjecture, we perform additional tests. Specifically, we examine the effect of ETF activity on short-run informational efficiency conditional on information environment. We use firm size and analyst following to capture the information environment. Small firms and firms with low analyst following have less publicly available information; hence, the information asymmetry among market participants may be substantial. Further, limits to arbitrage should be greater for small firms than for big firms, which would reduce informational efficiency for small firms. Therefore, if our conjecture is correct, we should observe greater increases in informational efficiency for small firms (firms with fewer analysts following) than for big firms (firms with more analysts following).

At the beginning of each quarter, we classify the full sample into big and small stocks using NYSE median breakpoints. We would expect the effect of ETF activity on the returns–earnings relation to be stronger for small stocks than for big stocks. To test this conjecture, we redo the Fama and MacBeth (1973) regression as in Equation (12) for big and small stocks.

We classify firms into categories of high and low analyst following based on the number of analysts following a firm during each quarter. A firm is classified as having high (low) analyst coverage if its number of analysts following is greater (less) than the 75th percentile. The number of analysts is right skewed for the cross-section of stocks. For example, during the first quarter of 2013, the median number of analysts is 5, whereas the mean is around 8. We, therefore, adopt the 75th percentile as the breakpoint—roughly 11 analysts during that quarter—which we believe is a more reasonable classification. We expect that the effect of ETF activity on short-run informational efficiency is stronger for firms with low analyst coverage than for those with high analyst coverage. To test this conjecture, we redo the Fama and MacBeth (1973) regression as in Equation (12) for these partitions.

Table 3 presents the evidence for different partitions. Consistent with expectations, ETF activity increases

Table 2. ETF Activity and Contemporaneous Return–Earnings Relation

Variable	1	2	3	4	5	6	7
Intercept	0.022	0.021	0.03**	0.021	0.060**	0.060**	0.055***
	(1.22)	(1.17)	(2.07)	(1.17)	(2.67)	(2.46)	(2.77)
$Earn_{i,t}$	0.240***	0.224***	0.297***	0.233***	0.323***	0.287***	0.314***
	(6.30)	(6.78)	(6.55)	(6.36)	(5.26)	(4.92)	(4.93)
$\Delta ETF_{i,t}$	0.022***	0.023***		0.023***	0.022***	0.021***	0.020***
	(2.92)	(3.03)		(3.01)	(4.44)	(4.59)	(5.09)
$\boldsymbol{Earn_{i,t} \times \Delta ETF_{i,t}}$	**0.110****	**0.117****		**0.124****	**0.131****	**0.143*****	**0.125***
	(2.25)	**(2.29)**		**(2.30)**	**(2.64)**	**(2.77)**	**(2.00)**
$Inst\_residual_{i,t}$		0.457***	0.460***	0.461***	0.461***	0.461***	0.461***
		(5.54)	(5.81)	(5.77)	(5.54)	(5.53)	(5.74)
$Earn_{i,t} \times Inst\_residual_{i,t}$			0.460	0.415	0.235	0.196	0.270
			(1.23)	(1.09)	(0.63)	(0.52)	(0.70)
$MTB_{i,t-1}$					−0.000	−0.000	0.000
					(−0.21)	(−0.14)	(0.07)
$Size_{i,t-1}$					−0.005***	−0.005**	−0.005**
					(−2.80)	(−2.62)	(−2.52)
$STD_{i,t-1}$					−1.345**	−1.383**	−1.240*
					(−2.28)	(−2.32)	(−1.98)
$Ret_{i,\ (t-12,t-2)}$					−0.027	−0.030	−0.030
					(−1.29)	(−1.43)	(−1.46)
$Loss_{i,t}$					−0.034***	−0.033***	−0.034***
					(−7.29)	(−7.44)	(−8.12)
$Earn_{i,t} \times Loss_{i,t}$					−0.171***	−0.169***	−0.180**
					(−2.87)	(−2.73)	(−2.55)
$Earn_{i,t-1}$						0.240***	0.243***
						(7.75)	(4.27)
$ETF_{i,t-1}$						−0.042	−0.061
						(−0.41)	(−0.54)
$Earn_{i,t-1} \times \Delta ETF_{i,t}$							−0.016
							(−0.20)
$\beta_{i,t-1}$							0.005
							(1.07)
$Ret_{i,t+1}$							−0.017
							(−1.34)
Adjusted R^2	0.014	0.034	0.032	0.035	0.070	0.079	0.103
Observations	78,984	78,984	78,984	78,984	78,984	78,984	78,984

Notes. This table presents associations between returns and earnings using Fama and MacBeth (1973) regressions of returns on contemporaneous earnings. The tabulated coefficient estimates are the time-series averages from the following cross-sectional regression: $Ret_{i,t} = b_{0,t} + b_{1,t}Earn_{i,t} + b_{2,t}\Delta ETF_{i,t} + b_{3,t}Earn_{i,t} \times \Delta ETF_{i,t} + Controls_{i,t} + \varepsilon_{i,t}$, where $Ret_{i,t}$ is the compounded 3-month return for quarter t. $Earn_{i,t}$ is the seasonally adjusted earnings for quarter t, deflated by the price at the beginning of quarter t. $\Delta ETF_{i,t}$ is the within-quarter decile rank of changes in ETF ownership, scaled to (0.1,1). All other variables are as defined in the appendix. The t-statistics with Newey–West (1987) correction for autocorrelation are reported in parentheses.

$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

short-run informational efficiency for small firms and firms with low analyst coverage. In contrast, we are unable to document such improvements for big firms and firms with high analyst coverage. Specifically, as documented in columns (1) and (2), ETF activity increases the returns–earnings relation for small firms, but not for big firms. Ben-David et al. (2018) find that ETF arbitrage activity increases intraday and daily volatility and conclude that these results are consistent with ETF arbitrage propagating nonfundamental shocks to the underlying stocks. However, this effect is significant only for big firms. Thus, our results, rather than contradicting Ben-David et al. (2018), complement our understanding about how ETF activity affects small firms and big firms differently. Similarly, as documented in columns (3) and (4), ETF activity increases the returns–earnings relation only for firms with low analyst coverage, not for firms with high analyst coverage. These results collectively suggest that ETF activity increases the returns–earnings relation for firms with weak information environments, but not for firms with strong information environments.

Table 3. ETF Activity and Contemporaneous Return–Earnings Relation for Alternative Partitions

Variable	Small 1	Big 2	Low 3	High 4
Intercept	0.145***	0.181***	0.058***	0.051
	(7.02)	(6.90)	(2.75)	(1.55)
$Earn_{i,t}$	0.330***	0.484**	0.349***	0.443**
	(4.01)	(2.15)	(3.99)	(2.67)
$\Delta ETF_{i,t}$	0.042***	−0.015***	0.030***	−0.008*
	(7.59)	(−2.77)	(7.33)	(−2.01)
$\boldsymbol{Earn_{i,t} \times \Delta ETF_{i,t}}$	**0.148****	**−0.189**	**0.165****	**−0.037**
	(2.52)	**(−0.62)**	**(2.21)**	**(−0.24)**
$Inst\_residual_{i,t}$	0.589***	0.232***	0.502***	0.446***
	(7.28)	(4.67)	(5.45)	(9.58)
$Earn_{i,t} \times Inst\_residual_{i,t}$	0.502	0.616	0.294	0.354
	(0.96)	(0.59)	(0.60)	(0.55)
$MTB_{i,t-1}$	−0.000	0.001	0.000	−0.000
	(−0.46)	(1.26)	(0.19)	(−0.74)
$Size_{i,t-1}$	−0.025***	−0.018***	−0.007**	−0.004
	(−6.42)	(−5.30)	(−2.48)	(−1.31)
$STD_{i,t-1}$	−1.373**	9.756***	−1.246*	0.146
	(−2.18)	(3.66)	(−1.95)	(0.08)
$Ret_{i,\ (t-12,t-2)}$	−0.032*	−0.018	−0.034	−0.023
	(−1.79)	(−0.81)	(−1.65)	(−1.00)
$Loss_{i,t}$	−0.040***	−0.008	−0.035***	−0.033***
	(−9.34)	(−1.19)	(−8.02)	(−5.24)
$Earn_{i,t} \times Loss_{i,t}$	−0.197**	−0.353***	−0.239**	−0.135*
	(−2.56)	(−2.92)	(−2.60)	(−1.74)
$Earn_{i,t-1}$	0.249***	0.355**	0.267***	0.199
	(4.38)	(2.40)	(5.14)	(1.47)
$ETF_{i,t-1}$	0.228*	−0.001	−0.037	−0.016
	(1.77)	(−0.01)	(−0.31)	(−0.07)
$Earn_{i,t-1} \times \Delta ETF_{i,t}$	−0.005	−0.177	−0.007	−0.024
	(−0.07)	(−0.81)	(−0.11)	(−0.12)
$\beta_{i,t-1}$	−0.015	−0.026	−0.021	−0.012
	(−1.26)	(−1.28)	(−1.63)	(−0.67)
$Ret_{i,t+1}$	0.006	0.018***	0.005	0.010
	(1.35)	(2.99)	(1.09)	(1.41)
Adjusted R^2	0.111	0.158	0.102	0.15
Observed	52,837	26,147	54,162	24,822

Notes. This table presents associations between returns and earnings using Fama and MacBeth (1973) regressions for alternative subgroups. Small (big) firms are those with market capitalization below (above) the 50th NYSE percentile. Low (high) analyst coverage firms are those with number of analysts following below (above) the 75th percentile. For each subgroup, the tabulated coefficient estimates are the time-series averages from the following cross-sectional regression: $Ret_{i,t} = b_{0,t} + b_{1,t}Earn_{i,t} + b_{2,t}\Delta ETF_{i,t} + b_{3,t}Earn_{i,t} \times \Delta ETF_{i,t} + Controls_{i,t} + \varepsilon_{i,t}$, where $Ret_{i,t}$ is the compounded 3-month return for quarter t. $Earn_{i,t}$ is the seasonally adjusted earnings for quarter t, deflated by the price at the beginning of quarter t. $\Delta ETF_{i,t}$ is the within-quarter decile rank of changes in ETF ownership, scaled to (0.1,1). All other variables are as defined in the appendix. The t-statistics with Newey–West correction for autocorrelation are reported in parentheses.

$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

In summary, results presented in Tables 2 and 3 suggest that ETF activity is associated with increases in the short-run informational efficiency of contemporaneous accounting information by reflecting information for a broader section of stocks in a timely manner. Further, short-run informational efficiency increases for firms with weak information environments. In contrast, we are unable to document such improvement for firms with strong information environments.

4.2. ETF Activity and Incorporation of Components of Earnings Information

In this section, we investigate the channel through which ETF activity increases the short-run informational efficiency of underlying stocks. Common information that affects a basket of securities should result in ETF trading, as our model suggests that traders cannot profitably trade the ETF on idiosyncratic firm-specific information. Therefore, if ETF activity results in an increase in short-run informational efficiency for the underlying stocks, then such improvement should be attributable to systematic information rather than idiosyncratic information. We test this conjecture in this section. We decompose $Earn_{i,t}$ into two components, systematic and firm-specific earnings news, and regress $Ret_{i,t}$ on these two components and their interactions with $\Delta ETF_{i,t}$. To do so, we estimate the following Fama and MacBeth (1973) specification:

$$\begin{aligned} Ret_{i,t} = {} & b_{0,t} + b_{1,t}\Delta ETF_{i,t} + b_{2,t}Earn\_Sys_{i,t} \\ & + b_{3,t}Earn\_Firm_{i,t} + b_{4,t}Earn\_Sys_{i,t} \times \Delta ETF_{i,t} \\ & + b_{5,t}Earn\_Firm_{i,t} \times \Delta ETF_{i,t} + b_{6,t}Inst\_residual_{i,t} \\ & + b_{7,t}Earn\_Sys_{i,t} \times Inst\_residual_{i,t} \\ & + b_{8,t}Earn\_Firm_{i,t} \times Inst\_residual_{i,t} + b_{9,t}MTB_{i,t-1} \\ & + b_{10,t}Size_{i,t-1} + b_{11,t}STD_{i,t-1} + b_{12,t}Ret_{i,(t-12,t-2)} \\ & + b_{13,t}Loss_{i,t} + b_{14,t}Earn_{i,t} \times Loss_{i,t} + b_{15,t}Earn_{i,t-1} \\ & + b_{16,t}ETF_{i,t-1} + b_{17,t}Earn_{i,t-1} \times \Delta ETF_{i,t} \\ & + b_{18,t}\beta_{i,t-1} + b_{19,t}Ret_{i,t-1} + \varepsilon_{i,t}, \end{aligned} \tag{14}$$

where $Earn\_Sys_{i,t}$ is the systematic earnings news, and $Earn\_Firm_{i,t}$ is firm-specific earnings news. All other variables are defined as above.

The systematic earnings news is calculated as the fitted value from the quarterly regression for each stock i:

$$Earn_{i,t} = \beta_{0,i} + \beta_{1,i}Earn\_mkt_t + \beta_{2,i}Earn\_ind_{i,t} + \varepsilon_{i,t}, \tag{15}$$

where $Earn\_mkt_t$ is the weighted average of seasonally adjusted earnings of all firms whose earnings information is available in Compustat, and $Earn\_ind_{i,t}$ is the weighted average of seasonally adjusted earnings of all firms in the same two-digit SIC code as firm i. Firm-specific earnings news is obtained as the residuals of regression (15).

Table 4 presents the results. We analyze the full sample, as well as partitions based on firm size and analyst coverage. The evidence from Table 4 suggests that the increase in short-run informational efficiency is attributable to contemporaneous systematic accounting information. Specifically, in column (1), the coefficient on the interaction of $Earn\_Sys_{i,t}$ and $\Delta ETF_{i,t}$ is positive and significant, implying that an increase in ETF activity pushes prices to reflect more systematic fundamental information for the full sample. The coefficient on the interaction of $Earn\_Firm_{i,t}$ and $\Delta ETF_{i,t}$ is insignificant, indicating that the ETF activity does not increase short-run informational efficiency related to the firm-specific earnings news. Results are also consistent for different partitions. Specifically, for the small firms and firms with low analyst coverage—columns (2) and (4)—the coefficient on $Earn\_Sys_{i,t} \times \Delta ETF_{i,t}$ is significant and positive at the 0.01 and 0.05 levels, indicating that prices reflect more systematic earnings for firms with weak information environments. Consistent with expectations, we do not find improvement in short-run informational efficiency for big firms and firms with high analyst coverage, even after we split the earnings into systematic and firm-level earnings news.

Overall, the evidence suggests that increases in short-run informational efficiency are attributable to the timely incorporation of systematic accounting information.

4.3. ETF Activity Decomposition

The cross-sectional variation in change in ETF ownership can result from the addition or deletion of a stock from an ETF, from the creation and redemption process, or from both. The transaction cost effect, which is a function of the number of ETFs holding the firm, should be identified if it is there. Note, however, that the reduction of shorting costs due to the increase in the supply of lendable shares could occur either through an increase in the number of ETFs or through an increase in the ownership of the firm by a given set of ETFs. To investigate the relative effect of these channels on the short-run informational efficiency of underlying securities, we decompose the change in ETF ownership into the addition or deletion effect and its orthogonal component. More specifically, we regress change in ETF ownership on change in the number of ETFs that each stock is part of, and we estimate the fitted and residual value:

$$\Delta ETF_{i,t} = \beta_{0,i} + \beta_{1,i}\Delta\#ETF_{i,t} + \varepsilon_{i,t}, \tag{16}$$

where $\Delta\#ETF_{i,t}$ is the change in number of ETFs that follow stock i. The fitted values from the above specification are our measure of the ETF activity that is attributable to addition and deletion ($\#ETF\_Fitted_{i,t}$), and the orthogonal component is our measure of the ETF activity that is attributable to the creation and redemption process ($\Delta ETF\_Residual_{i,t}$). To investigate the short-run informational efficiency attributable to these components, we reestimate specification (12) after replacing the main and interaction effects of

Table 4. ETF Activity and Contemporaneous Return–Earnings Components Relation

Variable	Full 1	Small 2	Big 3	Low 4	High 5
Intercept	0.055***	0.144***	0.179***	0.056**	0.052*
	(2.77)	(6.84)	(7.38)	(2.66)	(1.72)
$\Delta ETF_{i,t}$	0.020***	0.042***	−0.014**	0.031***	−0.006
	(5.05)	(7.03)	(−2.42)	(6.83)	(−1.32)
$Earn\_Sys_{i,t}$	0.361***	0.325***	0.649*	0.374***	0.579**
	(4.15)	(3.81)	(1.99)	(4.03)	(2.44)
$Earn\_Firm_{i,t}$	0.300***	0.337***	0.467*	0.353***	0.425***
	(5.05)	(4.43)	(1.74)	(4.36)	(3.20)
$Earn\_Sys_{i,t} \times \Delta ETF_{i,t}$	**0.330***	**0.425***	**0.082**	**0.359**	**0.229**
	(2.85)	**(3.39)**	**(0.18)**	**(2.26)**	**(0.94)**
$Earn\_Firm_{i,t} \times \Delta ETF_{i,t}$	**0.066**	**0.062**	**−0.211**	**0.079**	**−0.056**
	(0.95)	**(0.92)**	**(−0.63)**	**(1.21)**	**(−0.41)**
$Inst\_residual_{i,t}$	0.467***	0.593***	0.258***	0.503***	0.439***
	(6.37)	(8.04)	(5.00)	(5.93)	(9.50)
$Earn\_Sys_{i,t} \times Inst\_residual_{i,t}$	−0.023	−0.214	1.357	0.170	1.339
	(−0.04)	(−0.37)	(0.79)	(0.16)	(0.72)
$Earn\_Firm_{i,t} \times Inst\_residual_{i,t}$	0.353	0.780	0.049	0.424	0.102
	(0.66)	(0.97)	(0.04)	(0.64)	(0.11)
$MTB_{i,t-1}$	0.000	−0.000	0.001	0.000	−0.000
	(0.01)	(−0.62)	(1.07)	(0.10)	(−0.85)
$Size_{i,t-1}$	−0.005**	−0.025***	−0.018***	−0.007**	−0.004
	(−2.66)	(−6.72)	(−5.88)	(−2.61)	(−1.46)
$STD_{i,t-1}$	−1.282**	−1.396**	10.137***	−1.254**	−0.331
	(−2.14)	(−2.35)	(3.46)	(−2.09)	(−0.17)
$Ret_{i,\ (t-12,t-2)}$	−0.030	−0.033*	−0.019	−0.034*	−0.023
	(−1.49)	(−1.82)	(−0.86)	(−1.70)	(−1.02)
$Loss_{i,t}$	−0.034***	−0.040***	−0.007	−0.035***	−0.033***
	(−8.71)	(−9.91)	(−1.21)	(−8.18)	(−5.60)
$Earn_{i,t} \times Loss_{i,t}$	−0.197***	−0.215***	−0.379***	−0.255***	−0.160*
	(−2.93)	(−2.94)	(−3.15)	(−3.01)	(−1.81)
$Earn_{i,t-1}$	0.245***	0.254***	0.382**	0.261***	0.216
	(4.77)	(4.88)	(2.53)	(5.27)	(1.62)
$ETF_{i,t-1}$	−0.066	0.223	−0.015	−0.041	−0.024
	(−0.50)	(1.65)	(−0.08)	(−0.30)	(−0.11)
$Earn_{i,t-1} \times \Delta ETF_{i,t}$	−0.021	−0.015	−0.203	0.001	−0.084
	(−0.26)	(−0.19)	(−0.76)	(0.02)	(−0.40)
$\beta_{i,t-1}$	−0.018	−0.015	−0.030	−0.021	−0.015
	(−1.21)	(−1.16)	(−1.37)	(−1.48)	(−0.70)
$Ret_{i,t+1}$	0.005	0.006	0.019***	0.005	0.011
	(1.21)	(1.50)	(2.79)	(1.23)	(1.54)
Adjusted R^2	0.107	0.115	0.164	0.107	0.157
Observed	78,984	52,837	26,147	54,162	24,822

Notes. This table presents associations between returns and earnings components using Fama and MacBeth (1973) regressions for alternative subgroups. Small (big) firms are those with market capitalization below (above) the 50th NYSE percentile. Low (high) analyst coverage firms are those with number of analysts following below (above) the 75th percentile. For each firm in each quarter t, total earnings are decomposed into systematic earnings news ($Earn\_Sys_{i,t}$) and firm-specific earnings news ($Earn\_Firm_{i,t}$). The systematic earnings news and firm-specific earnings news are calculated as the fitted values and residuals from firm-specific quarterly regressions of earnings on value-weighted market-related earnings and value-weighted industry-related earnings. For each subgroup, the tabulated coefficient estimates are the time-series averages from the following cross-sectional regression: $Ret_{i,t} = b_{0,t} + b_{1,t}Earn.Sys_{i,t} + b_{2,t}Earn.Firm_{i,t} + b_{3,t}\Delta ETF_{i,t} + b_{4,t}Earn.Sys_{i,t} \times \Delta ETF_{i,t} + b_{5,t}Earn.Firm_{i,t} \times \Delta ETF_{i,t} + Controls_{i,t} + \varepsilon_{i,t}$. All other variables are as defined in the appendix. The t-statistics with Newey–West correction for autocorrelation are reported in parentheses.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

$\Delta ETF_{i,t}$ with the decomposed terms and their respective interactions with earnings:

$$\begin{aligned}Ret_{i,t} = {} & b_{0,t} + b_{1,t}\#ETF\_Fitted_{i,t} + b_{2,t}\Delta ETF\_Residual_{i,t}\\ & + b_{3,t}Earn_{i,t} + b_{4,t}Earn_{i,t} \times \#ETF\_Fitted_{i,t}\\ & + b_{5,t}Earn_{i,t} \times \Delta ETF\_Residual_{i,t}\\ & + b_{6,t}\Delta ETF\_Residual_{i,t} + b_{7,t}MTB_{i,t-1}\\ & + b_{8,t}Size_{i,t-1} + b_{9,t}Std_{i,t-1} + b_{10,t}Ret_{i,(t-12,t-2)}\\ & + b_{11,t}Loss_{i,t} + b_{12,t}Loss_{i,t} \times Earn_{i,t} + b_{13,t}ETF_{i,t-1}\\ & + b_{14,t}Earn_{i,t-1} \times \Delta ETF_{i,t} + b_{15,t}\beta_{i,t-1}\\ & + b_{16,t}Ret_{i,t+1} + \varepsilon_{i,t}, \end{aligned} \tag{17}$$

where $\#ETF\_Fitted_{i,t}$ is the fitted value from Equation (16) and $Resid\_\Delta ETF_{i,t}$ is the residual from Equation (16). All other variables are defined above. Table 5 presents the results. We find that both channels increase the short-run informational efficiency of underlying securities. Specifically, as documented in column (1) using the full sample, the fitted (residual) value of $\Delta ETF_{i,t}$ interacted with earnings has a coefficient estimate of 10.286 (9.271) and is statistically significant at the 5% (10%) level. Columns (2) to (5) of Table 5 presents results for the alternative partitions. Results from these partitions also suggest that both channels increase the short-run informational efficiency of underlying securities.

4.4. ETF Activity and Comovement

In this section, we examine the link between ETF activity and return comovement. The literature finds that index membership increases comovement, and this increase is driven by nonfundamental factors (Harris and Gurel 1986, Vijh 1994, Barberis et al. 2005, Peng and Xiong 2006, Da and Shive 2016). So far, however, the evidence from this paper suggests that ETF activity increases the short-run informational efficiency of underlying stocks by incorporating systematic accounting information into prices for a broader cross-section of stocks. Therefore, an increase in comovement could also be driven by timely incorporation of systematic accounting information (i.e., fundamental factors) into stocks.

We investigate the relation between ETF activity and comovement in two steps. We first estimate ETF activity that is due to fundamental information and then link quarterly change in comovement to the fundamental-related ETF activity. Specifically, we adopt CAPM beta, $\beta_{i,t}$, to capture comovement (Da and Shive 2016), where market beta is estimated as the coefficient from the regression of stock i's daily excess return on the market daily return in quarter t.

In step 1, we estimate ETF activity that is attributable to fundamental information. Specifically, we estimate the following Fama and MacBeth (1973) regression:

$$\begin{aligned}|\Delta ETF_{i,t}| = {} & b_{0,t} + b_{1,t}|Earn\_Sys_{i,t}| + b_{2,t}|Earn\_Firm_{i,t}|\\ & + b_{3,t}\beta_{i,t-1} + b_{4,t}MTB_{i,t-1} + b_{5,t}Size_{i,t-1}\\ & + b_{6,t}STD_{i,t-1} + b_{7,t}ETF_{i,t-1} + b_{8,t}Ret_{i,(t-12,t-2)}\\ & + b_{9,t}Inst\_residual_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{18}$$

where $|Earn\_Sys_{i,t}|$ is the absolute value of the systematic component of earnings, $|Earn\_Firm_{i,t}|$ is the absolute value of the firm-specific component of earnings, and $\beta_{i,t-1}$ is market beta at the beginning of quarter t. All other variables are as defined above.

In step 2, we examine what fraction of the increase in comovement is a result of ETF activity attributable to fundamental factors such as earnings. To do so, we regress quarterly change in market beta on the fitted value and residual from step 1 and other controls. Specifically, we estimate the following Fama and MacBeth (1973) regression:

$$\begin{aligned}\Delta\beta_{i,t} = {} & b_{0,t} + b_{1,t}Funda_{i,t} + b_{2,t}Other_{i,t} + b_{3,t}\beta_{i,t-1}\\ & + b_{4,t}Size_{i,t-1} + b_{5,t}MTB_{i,t-1} + b_{6,t}STD_{i,t-1}\\ & + b_{7,t}ETF_{i,t-1} + b_{8,t}Ret_{i,(t-12,t-2)}\\ & + b_{9,t}Inst\_Resid_{i,t-1} + \varepsilon_{i,t}, \end{aligned} \tag{19}$$

where $Funda_{i,t}$ and $Other_{i,t}$ are the fitted value and residual from Equation (18), respectively. All other variables are defined above. In step 1, we use the absolute value of earnings news because both buying and selling activity of ETFs due to fundamental information would push up return comovement. For example, facing positive economic news, traders would buy ETF shares and push stock prices to reflect positive news. Similarly, facing negative economic news, traders would sell ETF shares and push stock prices to reflect negative news. Both directions would increase stocks' return comovement. The absolute value of ETF ownership change captures the strength of trading, and more trading of ETFs in either direction should be positively associated with return comovement.

Table 6 (panels A and B) reports the results. Table 6, panel A presents results from step 1. A positive and significant coefficient on $|Earn\_Sys_{i,t}|$ in columns (2), (4), and (6) implies that more systematic fundamental information is associated with more ETF activity for firms with weak information environments.[12] Prior literature links ETF activity to nonfundamental factors, but our evidence suggests that ETF activity can be attributed at least partly to fundamental information. The second-step results are reported in Table 6, panel B. In column (1), we find that the increase in beta is also positively related to $Other_{i,t}$,

Table 5. ETF Activity Decomposition

Variable	Full 1	Small 2	Big 3	Low 4	High 5
Intercept	0.067***	0.165***	0.177***	0.074***	0.050
	(3.39)	(7.76)	(7.88)	(3.52)	(1.28)
$Earn_{i,t}$	0.401***	0.438***	0.385***	0.454***	0.381***
	(4.07)	(3.71)	(2.83)	(3.97)	(3.31)
$\#ETF\_Fitted_{i,t}$	−0.186	0.374	−0.601**	0.065	−1.424**
	(−0.80)	(1.31)	(−2.14)	(0.23)	(−2.15)
$\Delta ETF\_Residual_{i,t}$	2.536***	3.313***	0.866	2.765***	1.094**
	(6.30)	(7.80)	(1.42)	(7.42)	(2.03)
$\boldsymbol{Earn_{i,t} \times \#ETF\_Fitted_{i,t}}$	**10.286****	**12.089*****	**−29.788**	**16.027*****	**−21.330**
	(2.55)	**(3.36)**	**(−1.51)**	**(3.04)**	**(−1.19)**
$\boldsymbol{Earn_{i,t} \times \Delta ETF\_Residual_{i,t}}$	**9.271***	**10.606****	**−16.291**	**8.641**	**−7.108**
	(1.74)	**(2.18)**	**(−1.06)**	**(1.51)**	**(−0.38)**
$MTB_{i,t-1}$	0.000	−0.000	0.001	0.000	−0.000
	(0.06)	(−0.48)	(1.16)	(0.16)	(−0.72)
$Size_{i,t-1}$	−0.006**	−0.026***	−0.018***	−0.007**	−0.005
	(−2.68)	(−6.42)	(−6.17)	(−2.69)	(−1.61)
$STD_{i,t-1}$	−1.244**	−1.373**	8.992***	−1.240**	0.131
	(−2.09)	(−2.30)	(3.16)	(−2.06)	(0.08)
$Ret_{i,\ (t-12,\ t-2)}$	−0.033	−0.036*	−0.019	−0.037*	−0.024
	(−1.58)	(−1.94)	(−0.83)	(−1.77)	(−1.07)
$Earn_{i,t-1}$	0.244***	0.248***	0.331**	0.278***	0.151
	(4.19)	(4.24)	(2.18)	(4.83)	(1.08)
$ETF_{i,t-1}$	−0.100	0.233*	0.003	−0.067	−0.037
	(−0.87)	(1.74)	(0.01)	(−0.57)	(−0.16)
$Loss_{i,t}$	−0.033***	−0.038***	−0.008	−0.033***	−0.030***
	(−7.98)	(−9.14)	(−1.31)	(−8.15)	(−4.82)
$Earn_{i,t} \times Loss_{i,t}$	−0.183**	−0.200**	−0.391***	−0.246**	−0.110
	(−2.51)	(−2.52)	(−2.93)	(−2.65)	(−1.18)
$Inst\_residual_{i,t}$	0.450***	0.576***	0.217***	0.487***	0.458***
	(5.91)	(7.38)	(5.05)	(5.60)	(9.56)
$Earn_{i,t} \times Inst\_residual_{i,t}$	0.311	0.512	0.969	0.305	0.907
	(0.79)	(0.95)	(0.82)	(0.63)	(1.41)
$Earn_{i,t-1} \times \Delta ETF_{i,t}$	−0.033	−0.020	−0.170	−0.046	0.023
	(−0.42)	(−0.26)	(−0.68)	(−0.58)	(0.11)
$\beta_{i,t-1}$	0.006	0.007	0.018***	0.005	0.010
	(1.20)	(1.56)	(3.07)	(1.25)	(1.49)
$Ret_{i,t+1}$	−0.018	−0.016	−0.024	−0.021	−0.014
	(−1.37)	(−1.33)	(−1.29)	(−1.63)	(−0.77)
Adjusted R^2	0.110	0.119	0.168	0.110	0.160
Observed	78,984	52,837	26,147	54,162	24,822

Notes. This table presents associations between returns and earnings after decomposing ETF activity into the ETF addition/deletion effect and the orthogonal component. For each firm, ETF activity is decomposed into the ETF addition and deletion effect ($\#ETF\_Fitted_{i,t}$) and the orthogonal component ($\Delta ETF\_Residual_{i,t}$). These components are calculated as the fitted values ($\#ETF\_Fitted_{i,t}$) and residuals ($\Delta ETF\_Residual_{i,t}$) from the firm-specific regressions of ETF activity on quarterly change in number of ETFs that follow stock *i*. For each subgroup, the tabulated coefficient estimates are the time-series averages from the following cross-sectional regression: $Ret_{i,t} = b_{0,t} + b_{1,t}\#ETF\_Fitted_{i,t} + b_{2,t}\Delta ETF\_Residual_{i,t} + b_{3,t}Earn_{i,t} + b_{4,t}Earn_{i,t} \times \#ETF\_Fitted_{i,t} + b_{5,t}Earn_{i,t} \times \Delta ETF\_Residual_{i,t} + Controls + \varepsilon_{i,t}$. All other variables are as defined in the appendix. The *t*-statistics with Newey–West correction for autocorrelation are reported in parentheses.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

indicating that non-fundamental-driven ETF activity can also increase market beta, which is consistent with prior literature on index membership (Harris and Gurel 1986, Vijh 1994, Barberis et al. 2005, Peng and Xiong 2006, Da and Shive 2016). Interestingly, the positive and significant coefficient on $Funda_{i,t}$ in column (1) implies that the increase in beta can also be explained by fundamental earnings news, after we control for other firm characteristics. This evidence is consistent with our conjecture that ETF activity

Table 6. Market Beta and Fundamentals

Variable	Full 1	Small 2	Big 3	Low 4	High 5
			Panel A: First stage		
Intercept	0.368***	0.158***	0.463***	0.309***	0.558***
	(10.42)	(3.75)	(4.22)	(13.23)	(6.35)
$\lvert Earn\_Sys\rvert_{i,t}$	**0.103**	**0.149****	**−0.144**	**0.158****	**−0.162**
	(1.54)	**(2.56)**	**(−0.66)**	**(2.67)**	**(−0.97)**
$\lvert Earn\_Firm\rvert_{i,t}$	**0.020**	**0.057*****	**0.058**	**0.017**	**0.066**
	(0.81)	**(3.01)**	**(0.72)**	**(0.65)**	**(0.89)**
$\beta_{i,t-1}$	0.035***	0.029***	0.023*	0.037***	0.017
	(5.35)	(4.46)	(1.74)	(5.06)	(1.58)
$ETF_{i,t-1}$	6.278***	5.807***	5.028***	6.458***	4.399***
	(6.85)	(7.22)	(7.32)	(6.75)	(7.24)
$Size_{i,t-1}$	−0.002	0.036***	−0.012	0.006*	−0.019*
	(−0.41)	(4.29)	(−1.02)	(1.86)	(−1.99)
$Ret_{i,\ (t-12,t-2)}$	0.020***	0.013	0.018**	0.021**	0.011
	(3.03)	(1.40)	(2.31)	(2.65)	(1.54)
$MTB_{i,t-1}$	0.000	0.000	−0.000	0.000	0.000
	(0.48)	(0.21)	(−0.26)	(0.08)	(0.82)
$STD_{i,t-1}$	−2.245*	−0.903	−7.237**	−1.930*	5.213*
	(−1.87)	(−1.10)	(−2.42)	(−1.84)	(1.69)
$Inst\_residual_{i,t}$	0.026	0.026	−0.024	0.032	0.014
	(0.81)	(0.95)	(−0.53)	(1.01)	(0.33)
Adjusted R^2	0.276	0.310	0.264	0.281	0.287
			Panel B: Second stage		
Intercept	0.060	−0.263	0.671**	0.120	0.257
	(0.23)	(−0.98)	(2.03)	(0.46)	(0.59)
$Funda_{i,t}$	**1.510*****	**1.680*****	**0.970**	**1.112***	**2.243****
	(3.26)	**(3.48)**	**(1.45)**	**(1.93)**	**(2.44)**
$Other_{i,t}$	0.123**	0.134**	0.026*	0.145**	0.046
	(2.47)	(2.40)	(1.79)	(2.59)	(1.56)
$\beta_{i,t-1}$	−0.473***	−0.529***	−0.328***	−0.475***	−0.479***
	(−10.77)	(−11.66)	(−5.29)	(−10.76)	(−9.00)
$ETF_{i,t-1}$	−2.454	−2.934	−5.699	1.265	−7.476
	(−0.50)	(−0.56)	(−0.99)	(0.19)	(−0.82)
$Ret_{i,\ (t-12,t-2)}$	0.083**	0.050	0.221***	0.086**	0.093
	(2.15)	(1.26)	(3.36)	(2.55)	(1.43)
$MTB_{i,t-1}$	0.002	0.003*	−0.001	0.001	0.001
	(1.19)	(1.90)	(−0.58)	(1.11)	(0.40)
$Size_{i,t-1}$	−0.033	0.019	−0.092**	−0.019	−0.083*
	(−1.20)	(0.65)	(−2.14)	(−0.72)	(−1.82)
$STD_{i,t-1}$	−2.267	−0.875	−0.103	−3.434	−2.963
	(−0.58)	(−0.22)	(−0.01)	(−1.04)	(−0.30)
$Inst\_residual_{i,t}$	0.208**	0.216**	0.084	0.125	0.353*
	(2.11)	(2.18)	(0.45)	(1.20)	(1.92)
Adjusted R^2	0.288	0.298	0.286	0.289	0.309

Notes. Table 6 presents associations between increase in market $\beta(\beta_{i,t})$ and ETF activity linked to fundamentals. $\beta_{i,t}$ is the coefficient of the stock's daily excess returns on daily market excess returns in quarter t. $\Delta\beta_{i,t}$ is the difference between $\beta_{i,t}$ and $\beta_{i,t-1}$. Panel A presents the relation between ETF activity, systematic and firm-specific earnings news, and other controls. The tabulated coefficient estimates in Panel A are the time-series averages from the following cross-sectional regression: $|\Delta ETF_{i,t}| = b_{0,t} + b_{1,t}|Earn\_Sys_{i,t}| + b_{2,t}|Earn\_Firm_{i,t}| + Controls_{i,t} + \varepsilon_{i,t}$, where $|\Delta ETF_{i,t}|$ is the within-quarter decile rank of the absolute value of $\Delta ETF_{i,t}$, scaled to (0.1, 1). Panel B presents the relation between $\Delta\beta_{i,t}$ and fundamental-related and non-fundamental-related factors. The tabulated coefficient estimates in Panel B are the time-series averages from the following cross-sectional regression: $\Delta\beta_{i,t} = b_{0,t} + b_{1,t}Funda_{i,t} + b_{2,t}Other_{i,t} + Controls_{i,t} + \varepsilon_{i,t}$. $Funda_{i,t}$ is the fitted value of $|\Delta ETF_{i,t}|$ obtained from the first-stage regression (panel A), and $Other_{i,t}$ is the residual obtained from the first-stage regression. All other variables are as defined in the appendix. The t-statistics with Newey–West correction for autocorrelation are reported in parentheses.

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

incorporates systematic accounting information in a timely manner, resulting in increases in both short-run informational efficiency and return comovement. We find similar results for stocks with weak information environments in columns (2) and (4). In summary, we find evidence to support the claim that the increase in comovement can be partially explained by fundamental-driven ETF activity.

4.5. ETF Activity and Post-Earnings-Announcement Drift

In this section, we investigate the relation between ETF activity and postearnings-announcement drift. PEAD is a longstanding phenomenon in which prices drift in the direction of earnings surprises (e.g., Ball and Brown 1968 and Chordia and Shivakumar 2006). If ETF activity incorporates information into underlying stock prices in a timely manner, then we expect that PEAD should be attenuated with higher ETF activity. We start this analysis by confirming that the PEAD strategy has positive raw and factor adjusted returns in our sample period. Table 7, panel A documents the results. At the beginning of each month, all stocks in our sample are sorted into quintile portfolios based on their most recent $Earn_{i,t}$. The first quintile represents firms with the lowest $Earn_{i,t}$, and the fifth quintile represents firms with the highest $Earn_{i,t}$. In our sample period, we find that returns to the PEAD strategy are positive and significant. Specifically, we find that the PEAD strategy on average results in raw returns of 1.99% per month. Table 7, panel A also presents the factor-adjusted returns. Specifically, we employ the Fama and French (2015) model to calculate the risk-adjusted returns. Five-factor adjusted returns are 2.07% per month. To investigate the relation between the ETF activity and returns to the PEAD strategy, we do double independent sorts on five portfolios formed on ETF activity and standardized unexpected earnings ($Earn_{i,t}$). Raw returns and five-factor adjusted returns to these 25 portfolios are tabulated in Table 7, panel B. Consistent with expectations, we find that returns to the PEAD strategy attenuate as the ETF activity increases. Specifically, for the first quintile of ETF activity (low ETF activity), returns to the PEAD strategy are 2.53%. In contrast, for the fifth quintile of ETF activity (high ETF activity), returns to the PEAD strategy are 1.74%. The difference in the PEAD strategy returns between low and high ETF activity is 0.79%, and it is statistically significant at the 1% level. Table 7, panel B also presents the Fama and French (2015) five-factor adjusted returns. We find a similar pattern of results. In particular, the difference in the PEAD strategy five-factor adjusted returns between low and high ETF activity is 0.88%, and it is statistically significant at the 1% level. This evidence suggests that ETF activity incorporates information quickly into stock prices, resulting in lower PEAD.

Collectively, the evidence in Table 7, panels A and B, suggests that ETF activity incorporates information more quickly into stock prices, resulting in the attenuation of PEAD.

4.6. Russell 1000/2000 Index Reconstitution

In this section, we employ Russell 1000/2000 index reconstitution as an alternative setting to investigate the effects of ETF activity on the short-run informational efficiency of underlying securities. The Russell 1000 index includes the top 1,000 stocks ranked by market capitalization, whereas the Russell 2000 index includes the next 2,000 stocks. Each June, based on end-of-May stock market capitalization, the two indexes are reconstituted.[13] Chang et al. (2015) show that small market capitalization changes around rank 1,000 move stocks from one index to the other, but the firms close to this rank cutoff are similar in terms of firm characteristics, such as assets, earnings, and so on.[14] However, ETF ownership increases significantly for a stock that has a market capitulation rank close to 1,000 and moves from the Russell 1000 to the Russell 2000 index (Ben-David et al. 2018). More specifically, as documented by Chang et al. (2015), the weights of the largest stocks in the Russell 2000 are bigger than those of the smallest stocks in the Russell 1000 by a factor of 10, as the indexes are value weighted. Hence, when the smallest stock of the Russell 1000 index is switched to the Russell 2000 index, ETF ownership of the stock increases significantly.[15]

Therefore, Russell 1000/2000 index reconstitution can be used as a setting to test our conjecture that ETF activity increases short-run informational efficiency. Russell index constituents from 2004 to 2013 are from Russell Investments. One potential concern with this setting is that firms that move between the two indexes may be experiencing changes in market capitalization due to events such as mergers and acquisitions. To mitigate this concern, we use two data filters. First, we restrict our analysis to firms that move between indexes within the fixed bandwidth of ±100 (i.e., firms that are either among the bottom 100 firms of the Russell 1000 index, in terms of market capitalization rank, or among the top 100 firms of the Russell 2000 index). Second, we exclude the observations if the firm is engaged in a merger, takeover, or liquidation. We also include firm end-of-May CRSP market cap rankings *prior to* index reconstitution, as suggested by Appel et al. (2018). The final sample contains 272 firm-quarters that switch from the Russell 1000 to the Russell 2000, and 469 firm-quarters that switch from the Russell 2000 to the Russell 1000.[16]

Table 7. ETF Activity and Post-Earnings-Announcement Drift

Panel A: Unconditional sorting

$Earn_{i,t}$ rank	Return	5-factor alpha
1-Sell	–0.13%	–0.92%
2	0.37%	–0.38%
3	0.65%	0.02%
4	1.35%	0.64%
5-Buy	1.87%	1.15%
Buy–Sell	**1.99%*****	**2.07%*****

Panel B: Double sorts on unexpected earnings and ETF activity

	Return $Earn_{i,t}$ rank						5-factor alpha $Earn_{i,t}$ rank					
$\lvert\Delta ETF_{i,t}\rvert$ rank	1-Sell	2	3	4	5-Buy	Buy-Sell	1-Sell	2	3	4	5-Buy	Buy-Sell
1-Low	–0.34%	0.50%	0.86%	1.52%	2.19%	2.53%	–1.06%	–0.09%	0.18%	0.92%	1.60%	2.66%
2	–0.02%	0.34%	0.86%	1.37%	2.00%	2.02%	–0.91%	–0.51%	0.15%	0.56%	1.17%	2.08%
3	–0.04%	0.47%	0.74%	1.26%	1.66%	1.70%	–0.86%	–0.28%	0.04%	0.56%	0.99%	1.85%
4	0.13%	0.40%	0.73%	1.47%	2.05%	1.92%	–0.76%	–0.40%	0.04%	0.71%	1.12%	1.88%
5-High	–0.01%	0.34%	0.73%	1.17%	1.73%	1.74%	–0.91%	–0.38%	0.00%	0.35%	0.87%	1.78%
						0.79%***						**0.88%*****

Notes. Table 7 presents the results of the relation between ETF activity and postearnings-announcement drift. Standardized unexpected earnings ($Earn_{i,t}$) is calculated as the rolling seasonal random walk earnings, deflated by beginning of the quarter stock price. Panel A presents average monthly returns and Fama–French 5-factor adjusted returns for portfolios sorted by the most recent $Earn_{i,t}$ variable. At the beginning of each month, all the stocks are sorted into quintile portfolios based on their $Earn_{i,t}$. The first quintile represents firms with the lowest $Earn_{i,t}$, and the fifth quintile represents firms with the highest $Earn_{i,t}$. Also reported are returns on a Buy–Sell portfolio that longs $Earn_{i,t}$ Portfolio 5 and shorts $Earn_{i,t}$ Portfolio 1. In panel B, all stocks are independently sorted based on $|\Delta ETF_{i,t}|$ and $Earn_{i,t}$ into 5×5 portfolios, and we report the average monthly returns and Fama–French 5-factor adjusted returns for the 25 portfolios. Also reported are returns on the Buy–Sell portfolio that longs $Earn_{i,t}$ Portfolio 5 and shorts $Earn_{i,t}$ Portfolio 1, as well as the Buy–Sell portfolio returns for each ETF activity quintile.

***$p < 0.01$.

In Table 8, panel A, we confirm that firms transitioning from the Russell 1000 to the Russell 2000 experience an increase in the ETF ownership relative to firms transitioning from the Russell 2000 to the Russell 1000. Specifically, we conduct the following regression:

$$\begin{aligned} ETF_{i,t} = {} & b_0 + b_1 In_{i,t} \times Post_{i,t} + b_2 In_{i,t} + b_3 Post_{i,t} \\ & + b_4 CRSP\_MayRank_{i,t} + b_5 MTB_{i,t-1} \\ & + b_6 STD_{i,t-1} + b_7 Ret_{i,(t-12,t-2)} + b_8 Loss_{i,t} \\ & + b_9 Earn_{i,t} \times Loss_{i,t} + b_{10} Inst\_residual_{i,t} \\ & + b_{11} Earn_{i,t} \times Inst\_residual_{i,t} + b_{12} Earn_{i,t} \\ & + b_{13} Earn_{i,t-1} + b_{14} \beta_{i,t-1} + b_{15} Ret_{i,t+1} \\ & + \mu_t + \varepsilon_{i,t}, \end{aligned} \tag{20}$$

where $In_{i,t}$ equals 1 if firm i moves from the Russell 1000 to the Russell 2000 during the sample period and equals 0 if firm i moves from the Russell 2000 to the Russell 1000. $Post_{i,t}$ equals 1 for the reconstitution event quarter ahead and equals 0 for the quarter prior to the event date. $CRSP\_MayRank_{i,t}$ captures end-of-May CRSP market cap rankings. The regression also includes quarter fixed effects, μ_t, to control for time trends in short-run informational efficiency across all firms in our sample. The main variable of interest, b_1, captures the changes in ETF ownership for firms moving to the Russell 2000 relative to the changes for firms moving to the Russell 1000. A positive b_1 indicates that switching from the Russell 1000 to the Russell 2000 translates into increased ETF ownership.

Across the columns in Table 8, panel A, we document an increase in ETF ownership for a stock switching from the Russell 1000 to the Russell 2000, relative to a stock switching from the Russell 2000 to the Russell 1000, for the full sample, for small stocks, and for stocks with low analyst coverage.[17]

Given that the ETF ownership of a stock increases significantly if it switches from the Russell 1000 index to the Russell 2000 index, we expect stocks that move into the Russell 2000 index to have higher informational efficiency than stocks that move into the Russell 1000 index. As the ETF ownership of a stock that moves down from the Russell 1000 to the Russell 2000 increases significantly, the supply of lendable shares may increase dramatically as well. The stock that moves up to the Russell 1000 similarly should show a decline in lendable shares and a resulting decline in efficiency. Further, we should expect the difference in informational efficiency before and after the transition into the Russell 2000 index to be larger than the difference before and after the move into the Russell 1000 index. To test this conjecture, we employ

Table 8. ETF Activity and Return–Earnings Relation Around the Russell 1000/Russell 2000 Cutoff

Variable	Full 1	Small 2	Big 3	Low 4	High 5
Panel A: First stage					
$\boldsymbol{In_{i,t} \times Post_{i,t}}$	**0.012*****	**0.014*****	**0.007**	**0.013*****	**0.007**
	(4.01)	**(3.27)**	**(1.16)**	**(4.12)**	**(1.29)**
$In_{i,t}$	−0.011***	−0.012***	−0.011***	−0.011***	−0.004
	(−4.54)	(−3.49)	(−2.63)	(−3.98)	(−1.01)
$CRSP\_MayRank_{i,t-1}$	−0.000	−0.000**	0.000	−0.000	0.000
	(−0.58)	(−2.10)	(0.50)	(−1.28)	(0.63)
$MTB_{i,t-1}$	−0.000	0.000	−0.000	−0.000	−0.000
	(−0.82)	(0.26)	(−1.16)	(−0.71)	(−0.19)
$STD_{i,t-1}$	−0.836***	−0.628**	−0.959*	−0.976***	4.695*
	(−3.51)	(−2.58)	(−1.91)	(−3.96)	(1.75)
$Ret_{i,\ (t-12,\ t-2)}$	−0.000*	−0.000	−0.000	−0.001*	0.001
	(−1.80)	(−0.33)	(−1.18)	(−1.94)	(0.91)
$Loss_{i,t}$	−0.001	−0.000	−0.001	−0.001	0.002
	(−0.70)	(−0.05)	(−0.48)	(−0.70)	(0.52)
$Earn_{i,t} \times Loss_{i,t}$	−0.049*	−0.047	0.011	0.003	−0.231***
	(−1.65)	(−1.40)	(0.22)	(0.08)	(−3.30)
$Inst\_residual_{i,t}$	−0.013**	−0.019*	−0.012	−0.012	−0.024*
	(−2.17)	(−1.77)	(−1.54)	(−1.65)	(−1.73)
$Earn_{i,t} \times Inst\_residual_{i,t}$	−0.053	0.051	−0.680	0.050	−0.304
	(−0.58)	(0.56)	(−1.51)	(0.61)	(−0.48)
$Earn_{i,t}$	0.022	0.020	−0.027	−0.008	0.140***
	(0.85)	(0.71)	(−0.65)	(−0.31)	(3.27)
$Earn_{i,t-1}$	−0.003	−0.003	−0.014	−0.003	−0.023
	(−0.56)	(−0.45)	(−0.80)	(−0.45)	(−1.36)
$\beta_{i,t-1}$	−0.001	0.000	−0.004*	−0.001	−0.001
	(−0.83)	(0.04)	(−1.81)	(−0.65)	(−0.42)
$Ret_{i,t+1}$	0.006**	0.003	0.005	0.006*	0.007
	(2.04)	(0.73)	(1.56)	(1.87)	(1.39)
Intercept	0.048***	0.052***	0.048***	0.050***	0.030***
	(10.59)	(8.88)	(7.85)	(11.78)	(3.95)
Observed	741	318	423	572	169
Fixed effects	Quarter	Quarter	Quarter	Quarter	Quarter
Cluster	Firm	Firm	Firm	Firm	Firm
Adjusted R^2	0.691	0.643	0.732	0.692	0.748
Panel B: Second stage					
$\boldsymbol{Earn_{i,t} \times In_{i,t} \times Post_{i,t}}$	**5.557***	**15.298*****	**0.390**	**5.713***	**−0.824**
	(1.90)	**(4.63)**	**(0.28)**	**(1.86)**	**(−0.15)**
$Earn_{i,t}$	−0.128	2.186	1.024	−0.122	−0.121
	(−0.15)	(0.71)	(0.96)	(−0.12)	(−0.03)
$In_{i,t} \times Post_{i,t}$	0.165***	0.228**	0.134*	0.144***	0.293***
	(3.44)	(2.24)	(1.91)	(2.63)	(2.68)
$In_{i,t}$	−0.238***	−0.302***	−0.333***	−0.229***	−0.273***
	(−8.02)	(−4.77)	(−6.11)	(−6.65)	(−3.46)
$Earn_{i,t} \times In_{i,t}$	0.566	−2.631	2.170**	0.520	3.589
	(0.75)	(−0.84)	(1.98)	(0.62)	(0.83)
$Earn_{i,t} \times Post_{i,t}$	−4.884*	−15.544***	−1.053	−4.876*	−1.625
	(−1.75)	(−4.83)	(−0.92)	(−1.67)	(−0.30)
$CRSP\_MayRank_{i,t-1}$	0.000	−0.000	−0.000	0.000	−0.000
	(1.29)	(−0.90)	(−0.18)	(1.62)	(−0.52)
$MTB_{i,t-1}$	0.007**	0.006	0.001	0.008***	−0.004
	(2.48)	(1.15)	(0.52)	(2.69)	(−0.45)
$STD_{i,t-1}$	−2.684	−0.374	−1.487	−4.490	36.183
	(−0.53)	(−0.12)	(−0.22)	(−0.90)	(0.90)

Table 8. (Continued)

Variable	Full 1	Small 2	Big 3	Low 4	High 5
$Ret_{i,\ (t-12,\ t-2)}$	−0.003	0.009	−0.004*	−0.003	−0.003
	(−0.78)	(0.54)	(−1.66)	(−0.88)	(−0.13)
$Loss_{i,t}$	−0.022	−0.048	0.033	−0.036	0.060
	(−0.85)	(−1.53)	(1.09)	(−1.20)	(1.22)
$Earn_{i,t} \times Loss_{i,t}$	−1.501*	0.352	−2.567**	−1.755*	−1.364
	(−1.89)	(0.72)	(−2.21)	(−1.71)	(−0.95)
$Inst\_residual_{i,t}$	0.042	0.009	−0.039	0.077	−0.177*
	(0.86)	(0.10)	(−0.58)	(1.35)	(−1.68)
$Earn_{i,t} \times Inst\_residual_{i,t}$	0.026	0.028	0.003	0.041*	−0.045
	(1.29)	(0.93)	(0.20)	(1.81)	(−1.24)
$Earn_{i,t-1}$	0.908***	1.089***	0.381***	0.950***	0.718**
	(3.56)	(4.88)	(2.74)	(3.09)	(2.20)
$Earn_{i,t-1} \times In_{i,t}$	−7.960	0.002	3.108	−9.240*	8.598
	(−1.60)	(0.00)	(0.50)	(−1.75)	(1.61)
$Earn_{i,t-1} \times Post_{i,t}$	−0.040	1.427*	−0.653	−0.043	0.814
	(−0.05)	(1.74)	(−1.50)	(−0.05)	(0.41)
$Earn_{i,t-1} \times In_{i,t} \times Post_{i,t}$	0.566	−0.898	0.476	0.537	−0.786
	(0.67)	(−1.07)	(0.89)	(0.58)	(−0.40)
$\beta_{i,t-1}$	1.925*	10.079	1.834***	1.754	1.919
	(1.67)	(1.01)	(2.70)	(1.39)	(0.65)
$Ret_{i,t+1}$	−2.197*	−10.243	−0.071	−1.932	−1.151
	(−1.78)	(−1.02)	(−0.09)	(−1.47)	(−0.41)
Intercept	−0.048	0.197*	0.038	−0.104	0.165
	(−0.80)	(1.75)	(0.57)	(−1.55)	(1.20)
Observed	741	318	423	572	169
Fixed effects	Quarter	Quarter	Quarter	Quarter	Quarter
Cluster	Firm	Firm	Firm	Firm	Firm
Adjusted R^2	0.424	0.722	0.428	0.453	0.330

Notes. Table 8 presents associations between returns and earnings using the sample comprising firms that are added to the Russell 1000 from the Russell 2000 index or moved from the Russell 1000 to the Russell 2000 index. The sample is from March 2004 to December 2013. Small (big) firms are those with market capitalization below (above) the 50th NYSE percentile. Low (high) analyst coverage firms are those with number of analysts following below (above) the 75th percentile. In panel A, for each subgroup, we estimate the following regression: $ETF_{i,t} = b_1 In_{i,t} + b_2 Post_{i,t} + b_3\, In_{i,t} \times Post_{i,t} + Controls + \mu_t + \varepsilon_{t,t}$, where $In_{i,t}$ equals 1 if firm i is moved from the Russell 1000 to the Russell 2000 index during the sample period, and $In_{i,t}$ equals 0 if firm i is moved from the Russell 2000 to the Russell 1000 index. $Post_{i,t}$ equals 1 for the reconstitution event quarter and the quarter ahead, and equals 0 for the quarter prior to the event date. In panel B, for each subgroup, we estimate the following regression: $Ret_{i,t} = b_1 Earn_{i,t} + b_2 In_{i,t} + b_3 Post_{i,t} + b_4\, Earn_{i,t} \times\ In_{i,t} \times Post_{i,t} + b_5 In_{i,t} \times Post_{i,t} + b_6 Earn_{i,t} \times In_{i,t} + b_7 Earn_{i,t} \times Post_{i,t} + Controls + \mu_t + \varepsilon_{t,t}$. The regressions also include quarter fixed effects, and all other variables are as defined in the appendix. Standard errors clustered by firm are reported in parentheses. For brevity, the coefficient estimates on quarter fixed effects are not reported.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

a difference-in-differences design. Specifically, we adopt the following pooled regression specification:

$$\begin{aligned}Ret_{i,t} = {} & b_0 + b_1 Earn_{i,t} \times In_{i,t} \times Post_{i,t} + b_2 Earn_{i,t} + b_3 In_{i,t} \\ & + b_4 Post_{i,t} + b_5 In_{i,t} \times Post_{i,t} + b_6 Earn_{i,t} \times In_{i,t} \\ & + b_7 Earn_{i,t} \times Post_{i,t} + b_8 CRSP\_MayRank_{i,t} \\ & + b_9 MTB_{i,t-1} + b_{10} STD_{i,t-1} + b_{11} Ret_{i,(t-12,t-2)} \\ & + b_{12} Loss_{i,t} + b_{13} Earn_{i,t} \times Loss_{i,t} + b_{14} Inst\_residual_{i,t} \\ & + b_{15} Earn_{i,t} \times Inst\_residual_{i,t} + b_{16} Earn_{i,t-1} \\ & + b_{17} Earn_{i,t-1} \times In_{i,t} + b_{18} Earn_{i,t-1} \times Post_{i,t} \\ & + b_{19} Earn_{i,t-1} \times In_{i,t} \times Post_{i,t} + b_{14} \beta_{i,t-1} + b_{15} Ret_{i,t+1} \\ & + \mu_t + \varepsilon_{i,t}, \end{aligned} \tag{21}$$

where $In_{i,t}$ equals 1 if firm i moves from the Russell 1000 to the Russell 2000 during the sample period, and equals 0 if firm i moves from the Russell 2000 to the Russell 1000. $Post_{i,t}$ equals 1 for the reconstitution event quarter ahead and equals 0 for the quarter prior to the event date. The regression also includes quarter fixed effects, μ_t, to control for time trends in short-run informational efficiency across all firms in our sample. The main variable of interest, b_1, captures the changes in short-run informational efficiency for firms moving to the Russell 2000 relative to the changes for firms moving to the Russell 1000. A positive b_1 indicates that short-run informational efficiency increases for the former relative to the latter.[18]

Table 8, panel B presents the results. The results for the full sample are reported in column (1). For the full sample, we find that the change (postevent period – preevent period returns–earnings relation) in short-run informational efficiency for firms moving to the Russell 2000 is statistically greater than the change in short-run informational efficiency for firms moving to the Russell 1000. To partition firms into small versus big, for this analysis, we use the sample of addition and deletion firms rather than the full CRSP universe. If we use the CRSP universe, we do not find any small firms (firms with market capitalization below the NYSE 50th percentile) that are added to the Russell 2000 index. Similarly, we use the 75th percentile partition within the sample to classify firms into low and high analyst following partitions. Columns (2) to (5) present results for the alternative partitions. Consistent with expectations, short-run informational efficiency increases for firms with weaker information environments after they move to the Russell 2000 relative to firms moving to the Russell 1000, but not for firms with stronger information environments.[19]

Although we control for several observable firm characteristics in our research design, we cannot fully rule out the concurrent changes in other firm policies and institutional environment that could also potentially affect the short-run informational efficiency. This might pose a limitation in drawing inference using this setting. Notwithstanding the above limitation, the evidence using the Russell 1000/2000 index reconstitution setting corroborates our main finding of the impact of ETF activity on short-run informational efficiency. We find that ETF activity increases short-run informational efficiency for firms with weak information environments.

4.8. ETF Activity, Future Earnings, and Returns

Next, we investigate whether ETF activity incorporates current-quarter earnings news into current-quarter stock returns rather than whether long-term earnings news (e.g., one-year-ahead or two-year-ahead earnings news) is incorporated into stock prices on a timely basis. We do so because we expect that ETF activity does not predict future fundamental news; rather, it merely incorporates news from different information sources into stock prices in a timely manner. To validate this conjecture, we investigate whether ETF activity predicts one-quarter-ahead earnings news or stock returns. Specifically, we adopt the following Fama–MacBeth regression:

$$\begin{aligned} Earn(Ret)_{i,t+1} = {} & b_{0,t} + b_{1,t}\Delta ETF_{i,t} + b_{2,t}Ret_{i,t} + b_{3,t}Earn_{i,t} \\ & + b_{4,t}Earn_{i,t-1} + b_{5,t}Loss_{i,t} \\ & + b_{6,t}Loss_{i,t} \times Earn_{i,t} + b_{8,t}Size_{i,t} \\ & + b_{9,t}MTB_{i,t} + b_{10,t}STD_{i,t} + b_{11,t}ETF_{i,t-1} \\ & + b_{12,t}Ret_{i,(t-12,t-2)} + b_{12,t}Inst\_residual_{i,t} \\ & + \varepsilon_{i,t+1}, \end{aligned} \tag{22}$$

where $Earn_{i,t+1}$ is one-quarter-ahead seasonally adjusted earnings for firm i, and $Ret_{i,t+1}$ is one-quarter-ahead stock return for firm i. A positive $b_{1,t}$ indicates that ETF activity predicts future earnings information or future returns. The results are reported in Table 9. The evidence from all the specifications suggests that ETF activity is not significantly associated with one-quarter-ahead earnings or stock returns. Overall, the evidence does not suggest that ETF predicts either earnings news or stock returns.

4.9. ETF Activity and Active Share Lending

ETF activity might change the short-run informational efficiency of underlying securities, as some ETFs engage in share lending to short sellers. Therefore, if a stock is added to an ETF, shares available for shorting go up, thereby potentially reducing the cost of shorting those shares. We test the relation between the ETF activity and share lending using the Markit Securities Finance data set. *Share_Lending* is defined as a firm's active lendable shares divided by its shares outstanding, where active lendable shares are defined by Markit as the lendable quantity adjusted to remove lendable stock inventory that are not actively being made available for lending.[20]

To investigate the relation between ETF activity and share lending, we perform both portfolio analysis and regression analysis. Table 10, panel A documents the percentage of active share lending for portfolios sorted on ETF activity. We find that for the first decile of ETF activity (low ETF activity), on average active share lending increases by 0.34% during the quarter, whereas for the last decile of ETF activity (high ETF activity), active share lending increases by 1.07%. The difference is statistically significant. Multivariate regression results are reported in Table 10, panel B.

Table 9. ETF Activity, Future Earnings, and Returns

Variable	Quarter-ahead earnings				Quarter-ahead return			
Intercept	−0.001	−0.001	−0.000	−0.001	0.034*	0.033*	0.039**	0.057**
	(−0.45)	(−0.55)	(−0.40)	(−0.39)	(1.76)	(1.78)	(2.36)	(2.25)
$\Delta ETF_{i,t}$	**0.000**	**−0.001**	**−0.001**	**0.000**	**−0.004**	**−0.005**	**−0.006**	**−0.007**
	(0.20)	**(−0.84)**	**(−0.62)**	**(0.26)**	**(−0.61)**	**(−0.80)**	**(−1.01)**	**(−1.10)**
$Ret_{i,t}$		0.032***	0.027***	0.026***		0.002	−0.012	−0.014
		(6.37)	(6.01)	(6.20)		(0.11)	(−0.95)	(−0.89)
$Earn_{i,t}$			0.121***	0.121***			0.374***	0.371***
			(5.10)	(5.25)			(6.23)	(6.25)
$Earn_{i,t-1}$			0.055***	0.055***			−0.045	−0.033
			(3.82)	(3.92)			(−1.52)	(−1.63)
$Loss_{i,t}$			−0.002	−0.002*			−0.036***	−0.040***
			(−1.66)	(−1.90)			(−5.95)	(−8.89)
$Earn_{i,t} \times Loss_{i,t}$			0.022	0.021			−0.139**	−0.132*
			(0.94)	(0.89)			(−2.04)	(−1.97)
$Size_{i,t}$				−0.000				−0.003*
				(−0.75)				(−2.03)
$MTB_{i,t}$				0.000				0.000
				(0.84)				(0.44)
$STD_{i,t}$				0.176				−0.986**
				(0.91)				(−2.49)
$ETF_{i,t-1}$				0.009				−0.055
				(0.57)				(−0.57)
$Ret_{i,\ (t-12,t-2)}$				0.000				−0.021
				(0.22)				(−0.98)
$Inst\_residual_{i,t}$		0.005	0.003	0.003		−0.033	−0.032	−0.024
		(0.54)	(0.39)	(0.38)		(−0.89)	(−0.92)	(−0.96)
Adjusted R^2	0.000	0.012	0.058	0.061	0.002	0.015	0.044	0.063
Observed	78,936	78,936	78,936	78,936	78,936	78,936	78,936	78,936

Notes. Table 9 presents associations between one-quarter-ahead earnings (returns) and ETF activity using Fama and MacBeth (1973) regressions. The tabulated coefficient estimates are the time-series averages from the following cross-sectional regression: $Earn(Ret)_{i,t+1} = b_{0,t} + b_{1,t}\Delta ETF_{i,t} + Controls_{i,t} + \varepsilon_{i,t+1}$, where $Earn_{i,t+1}$ is the seasonally adjusted earnings for quarter $t + 1$, deflated by the price at the beginning of quarter $t + 1$. $Ret_{i,t+1}$ is the compounded 3-month return for quarter $t + 1$. $\Delta ETF_{i,t}$ is the within-quarter decile rank of changes in ETF ownership, scaled to (0.1,1). All other variables are as defined in the appendix. The t-statistics with Newey–West correction for autocorrelation are reported in parentheses.

$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Consistent with the portfolio sort results, we find that an increase in ETF activity is significantly associated with an increase in share lending. Overall, we find evidence consistent with our conjecture that ETF activity is positively associated with share lending.

4.10. Additional Analysis and Robustness Checks

We perform several additional analyses and robustness checks of our findings. First, our sample period includes the Great Recession period (Q4:2007–Q1:2009). To mitigate any concerns that the results might be driven by the crisis period, we drop the observations from Q4:2007–Q1:2009. We then repeat the main analyses. All of our main takeaways are robust to excluding the Great Recession period. Second, our main tests use within-quarter ranked ΔETF to proxy for ETF activity to mitigate outliers and to facilitate interpretation of the coefficient estimates. As a robustness check, we redo the main analysis using the raw ΔETF (instead of the rank variable). All main findings are robust to using this alternative measure. Finally, we also redo the analysis after adding one-quarter-ahead earnings and the interaction of ETF activity and one-quarter-ahead earnings to specification (12). We do not find that ETF activity incrementally incorporates future earnings into current stock returns.

5. Conclusions

We find that greater ETF activity is associated with improvement in short-run informational efficiency for underlying stocks. Further, we document that the increase in short-run informational efficiency is attributable to the incorporation of incrementally more systematic fundamental information, rather than firm-specific fundamental information. However, the improved short-run informational efficiency is confined to firms with weak information environments.

Prior literature finds that index membership increases comovement, and this increase is driven by

Table 10. ETF Activity and Share Lending

Panel A: Portfolio analysis

$\Delta ETF_{i,t}$ Rank	$\Delta Share\_Lending_{i,t}$
1-Low	0.34%
2	0.48%
3	0.46%
4	0.56%
5	0.58%
6	0.59%
7	0.62%
8	0.70%
9	0.83%
10-High	1.07%
High–Low	**0.73%*** **

Panel B: Regression analysis

Variable	Full 1	Small 2	Big 3	Low 4	High 5
Intercept	0.000	−0.005	0.003	−0.000	0.008
	(0.10)	(−1.42)	(0.72)	(−0.11)	(1.67)
$\boldsymbol{\Delta ETF_{i,t}}$	**0.007*****	**0.008*****	**0.007****	**0.008*****	**0.006*****
	(5.77)	**(4.75)**	**(2.16)**	**(4.85)**	**(3.15)**
$\Delta Inst\_resid_{,i,t}$	0.070***	0.078***	0.055***	0.079***	0.054***
	(5.55)	(6.31)	(5.09)	(5.37)	(8.50)
$MTB_{i,t-1}$	0.000	−0.000	0.000**	−0.000	0.000*
	(0.30)	(−0.08)	(2.06)	(−0.32)	(1.78)
$Size_{i,t-1}$	0.000	0.001***	−0.000	0.000	−0.001**
	(0.36)	(2.93)	(−1.00)	(0.97)	(−2.58)
$STD_{i,t-1}$	−0.002	0.018	−0.007	0.011	0.314
	(−0.04)	(0.34)	(−0.02)	(0.20)	(1.28)
$Ret_{i,\ (t-12,t-2)}$	0.004***	0.004***	0.001**	0.004***	0.003***
	(4.62)	(4.11)	(2.29)	(4.90)	(3.18)
$\beta_{i,t-1}$	−0.001***	−0.002***	−0.000	−0.001***	−0.001
	(−4.66)	(−4.48)	(−0.32)	(−5.65)	(−1.68)
$ETF_{i,t-1}$	0.099	0.086	0.067	0.088	0.086
	(1.37)	(1.25)	(0.74)	(1.21)	(1.11)
Adjusted R^2	0.064	0.077	0.038	0.077	0.039
Observed	69,881	46,950	22,931	27,696	22,185

Notes. Table 10 presents associations between ETF activity and change in share lending activity. Share lending activity ($Share\_Lending_{i,t}$) is defined as a firm's active lendable shares divided by its shares outstanding, where active lendable shares are defined as the lendable quantity adjusted to remove the lendable quantity that is not actively being made available for lending (data are from the Markit Securities Finance data set). Panel A reports average quarterly change in $\Delta\ Share\_Lending_{i,t}$ for 10 portfolios sorted by quarterly ETF activity ($\Delta ETF_{i,t}$). Each quarter, all the stocks are sorted into decile portfolios based on their $\Delta ETF_{i,t}$. The 1st quintile represents firms with the lowest $\Delta ETF_{i,t}$, and the 10th quintile represents firms with the highest $\Delta ETF_{i,t}$. Also reported are the differences between decile 10 and decile 1. In panel B, we report the regression analysis. Specifically, the tabulated coefficient estimates are the time-series averages from the following cross-sectional regression: $\Delta Share\_Lending_{i,t} = b_{0,t} + b_{1,t}\Delta ETF_{i,t} + Controls_{i,t} + \varepsilon_{i,t}$. $\Delta ETF_{i,t}$ is the within-quarter decile rank of changes in ETF ownership, scaled to (0.1,1). All other variables are as defined in the appendix. We report the regression results for the full sample and alternative subgroups. Small (big) firms are those with market capitalization below (above) the 50th NYSE percentile. Low (high) analyst coverage firms are those with number of analysts following below (above) the 75th percentile. The t-statistics with Newey–West correction for autocorrelation are reported in parentheses.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

nonfundamental factors (Harris and Gurel 1986, Vijh 1994, Barberis et al. 2005, Peng and Xiong 2006, Da and Shive 2016). However, if ETF activity is associated with more systematic information being incorporated into prices in a timely manner, that could increase return comovement. Consistent with expectations, we find that the increase in comovement linked to ETF activity can be partially explained by systematic fundamental information. Further, ETF activity is associated with attenuation of post-earnings-announcement drift and increase in active share lending. Finally, we corroborate our main findings using Russell 1000/2000 reconstitution as a setting. Collectively, the evidence presented in this paper suggests that ETF activity can improve short-run informational efficiency for underlying stocks.

Acknowledgments

The authors thank Dan Amiram, Francesco Franzoni (discussant), Jonathan Glover, Trevor Harris, Alon Kalay, Sharon Katz, Urooj Khan, Feng Li (discussant), Nan Li, Ross Lu (discussant), Matthew Lyle (discussant), Harry Mamaysky, Doron Nissim, Maria Ogneva, Stephen Penman, Shiva Rajgopal, Gil Sadka, Tano Santos, Suraj Srinivasan (editor), Mohan Venkatachalam, Russell Wermers, Ronghuo Zheng, two anonymous referees, and participants of the research seminars at Vanderbilt University, University of British Columbia, University of Maryland, the brown bag seminar at Columbia University, the American Accounting Association annual meeting, the MIT Asia Conference in Accounting, Western Finance Association meetings, and the London Trans-Atlantic Doctoral Conference for thoughtful discussions and suggestions. The authors thank Jason Chao of Russell Investments for providing Index membership data and Mohammadreza Bolandnazar for excellent research assistance. A previous version of this manuscript was titled "ETF Trading and Informational Efficiency of Underlying Securities." Any remaining errors are our own.

Appendix: Variable Definitions

Table A.1. ETF Variables

Variable	Description
$ETF_{i,t}$	The percentage of common shares outstanding of stock i held by ETFs at the end of quarter t. The data are obtained from Thomson–Reuters and CRSP.
$\Delta ETF_{i,t}$	At the end of each quarter t, change in $ETF_{i,t}$ from the end of quarter $t-1$ through the end of quarter t is ranked from 1 to 10 and scaled by 10. The data are obtained from Thomson–Reuters and CRSP.
$\lvert\Delta ETF_t\rvert$	At the end of each quarter t, the absolute value of change in $ETF_{i,t}$ from the end of quarter $t-1$ through the end of quarter t is ranked from 1 to 10 and scaled by 10. The data are obtained from Thomson–Reuters and CRSP.
$\#ETF_{i,t}$	The number of ETFs holding stock i at the end of quarter t. The data are obtained from Thomson–Reuters.

Table A.2. Main Variables

Variable	Description
$Ret_{i,t}$	The quarterly return for stock i in quarter t. The data are obtained from CRSP.
$Earn_{i,t}$	Seasonally adjusted earnings innovation, scaled by price: $Earn_{i,t} = \frac{X_{i,t} - X_{i,t-4}}{P_{i,t-1}}$, where $X_{i,t}$ denotes firm i's earnings per share excluding extraordinary items in quarter t, and $P_{i,t-1}$ denotes the price at the beginning of quarter t. The data are obtained from COMPUSTAT.
$Earn\_Sys_{i,t}$	Systematic component of the current earnings innovation. It is calculated as the fitted value from the quarterly regression for stock i: $Earn_{i,t} = \beta_{0,i} + \beta_{1,i}Earn\_mkt_t + \beta_{2,i}Earn\_ind_{i,t} + \varepsilon_{i,t}$, where $Earn\_mkt_t$ is the weighted average of seasonally adjusted earnings of all firms with available earnings information in Compustat, and $Earn\_ind_{i,t}$ is the weighted average of seasonally adjusted earnings of all firms in the same two-digit SIC code as firm i. The data are obtained from COMPUSTAT.
$Earn\_Firm_{i,t}$	Firm-specific component of the current earnings innovation. It is calculated as the residual from the quarterly regression for stock i: $Earn_{i,t} = \beta_{0,i} + \beta_{1,i}Earn\_mkt_t + \beta_{2,i}Earn\_ind_{i,t} + \varepsilon_{i,t}$, where $Earn\_mkt_t$ is the weighted average of seasonally adjusted earnings of all firms with available earnings information in Compustat, and $Earn\_ind_{i,t}$ is the weighted average of seasonally adjusted earnings of all firms in the same two-digit SIC code as firm i. The data are obtained from COMPUSTAT.
$\beta_{i,t}$	The $\beta_{i,t}$ coefficient from firm-quarter estimation of the model: $RET\_RF_{i,d} = \beta_{0,i} + \beta_{1,i}MKT\_RF_d + \varepsilon_{i,d}$, where $RET\_RF_{i,d}$ is the stock's excess return on day d, and MKT_RF_d is market excess return on day d. The model is estimated using daily returns in quarter t. The data are obtained from CRSP.

Table A.3. Other Variables

Variable	Description
$MVE_{i,t}$	The market value of equity at the end of quarter t. The data are obtained from CRSP.
$Size_{i,t}$	The logged market capitalization of the stock (in \$ millions) at the end of quarter t. The data are obtained from CRSP.
$BVE_{i,t}$	Book value of equity at the end of quarter t. The data are obtained from item #60 from COMPUSTAT.
$MTB_{i,t}$	The ratio of $MVE_{i,t}$ to $BVE_{i,t}$. The data are obtained from CRSP and COMPUSTAT.
$\#Analyst_{i,t}$	The number of analysts that follow firm i during quarter t. The data are obtained from the Institutional Broker's Estimate System.
$Inst\_residual_{i,t}$	Orthogonalized institutional ownership. It is calculated as the residual from the cross-sectional regression in each quarter: quarterly cross-$\Delta Inst_{i,t} = b_{0,t} + b_{1,t}\Delta ETF_{i,t} + \varepsilon_{i,t}$, where $\Delta Inst_{i,t}$ is the quarterly change in institutional ownership for stock i in quarter t. The data are obtained from the Thomson–Reuters s34 database.
$STD_{i,t}$	Standard deviation of firm i's earnings per share excluding extraordinary items, over 20 quarters prior to quarter t. The data are obtained from COMPUSTAT.
$Loss_{i,t}$	Indicator variable equaling 1 if $Earn_{i,t}$ is negative, and equaling 0 otherwise. The data are obtained from COMPUSTAT.
$Ret_{i,(t-12,t-2)}$	Compounded stock return between 12 months and 2 months prior to quarter t. The data are obtained from CRSP.
$Share\_Lending_{i,t}$	A firm's active lendable shares divided by its shares outstanding, where lendable shares are defined by Markit as the lendable quantity adjusted to remove the lendable stock inventory that is not actively being made available for lending. The data are obtained from the Markit Securities Finance data set.

Endnotes

[1] For example, the assets under management by ETFs grew from a total value of $417 billion in 2005 to $4.4 trillion as of September 2017 (Ernst & Young 2017). In 2017, the demand for domestic equity ETFs resulted in $186 billion net share issuances, whereas domestic equity mutual funds had net redemptions of $236 billion (Investment Company Institute 2018).

[2] Please see Piwowar (2017) for details.

[3] We find that ETF activity is positively related to active share lending. However, it should be noted that some ETFs—SPY and QQQ are important examples—are arranged as investment trusts and as such are not permitted to engage in share lending.

[4] To ensure that our results are not sensitive to the cutoff used to define the information environment, we repeat the analysis using alternative cutoffs. Specifically, we use the 66th percentile, 70th percentile, 80th percentile, and 90th percentile as alternative cutoffs for partitioning the sample. Our results are robust to these alternative cutoffs. We continue to find that ETF activity increases informational efficiency for stocks with weak information environments.

[5] See the *Economist* article available at http://www.economist.com/news/finance-and-economics/21627717-regulators-are-worried-trendy-new-product-will-sow-instability-emerging (accessed October 25, 2014).

[6] We believe this observation—that a Kyle model with multiple securities and a redundant security in which prices depend upon order flows from all securities does not have an equilibrium—to be new.

[7] There is a subtler reason to think that idiosyncratic information will not be revealed by expansion or contraction of ETF holdings. If idiosyncratically informed trade the ETF, in any given quarter, half of them will be buying and half selling, leaving the size of the ETF probabilistically unaffected. Further, idiosyncratically informed may buy the underlying and short the ETF to hedge out the systematic risk. Here again, probabilistically, there will be no effect on the ETF.

[8] We double-check our sample of ETFs using the CRSP mutual fund database. Specifically, we include an ETF in our sample only if *etflag* is equal to "F" in the CRSP mutual fund database.

[9] We start our analyses from 2004, because the average ETF ownership is below 1% before 2004.

[10] Our institutional ownership measure includes change in index mutual fund ownership. As a robustness check, we explicitly control for change in index mutual fund ownership in our specification (12). All our inferences remain unchanged.

[11] In untabulated analysis, we investigate whether the change in institutional ownership increases the informational efficiency of the underlying securities. Specifically, we reestimate specification (12) after replacing $Resid_{i,t}$ with $\Delta Inst_{i,t}$. We do not find robust evidence that change in institutional ownership increases the informational efficiency of underlying securities. All our inferences are robust to including either $Resid_{i,t}$ or $\Delta Inst_{i,t}$ in the specification.

[12] Although we focus on earnings information, ETF activity could incorporate common factor fundamental information, thereby increasing the return comovement.

[13] Chang et al. (2015) reject the hypothesis that firms have precise control over market capitalization; therefore, it seems unlikely that firms that are close to the index cutoff can manipulate their market capitalization.

[14] A number of studies use this setting: Chang et al. (2015), Boone and White (2015), Crane et al. (2016), Appel et al. (2016), and Chen et al. (2019).

[15] One of the authors attended the Financial Market Research Center's (Vanderbilt) conference on exchange products (5/18/2018). Several market participants noted that a Russell 2000 ETF is more popular than a Russell 1000 ETF because it is not approximated by an S&P 500 ETF.

[16] The main results are qualitatively similar if we use bandwidth = 150, 200, 250, or 500, or with no bandwidth.

[17] Consistent with the findings in Chang et al. (2015), we do not find that a transition from the Russell 1000 into the Russell 2000 has any effect on the overall institutional ownership.

[18] Importantly, we control for institutional ownership and the interaction between institutional ownership and earnings in all the regressions, in order to mitigate the concern that the results might be driven by other institutions (e.g., active institutions).

[19] If we use the CRSP universe to classify firms with high or low analyst following, the results are stronger for firms with low analyst following than the results presented in Table 8, panels A and B.

[20] In our main analyses, we use *Active Lendable Quantity* as defined by Markit as "Lendable quantity adjusted to remove lendable which is not being actively made available for lending." As a robustness check, we also redo the analyses with the variable *Lendable Quantity* as defined by Markit as "Quantity of stock inventory, available to lend." The results are qualitatively similar.

References

Abner D (2010) *The ETF Handbook* (John Wiley & Sons, Hoboken, NJ).

Antoniou C, Subrahmanyam A, Tosun O (2018) ETF ownership and corporate investment. Working paper, University of Warwick, Coventry, UK.

Appel I, Gormley T, Keim D (2016) Passive investors, not passive owners. *J. Financial Econom.* 121(1):111–141.

Appel I, Gormley T, Keim D (2018) Identification using Russell 1000/2000 index assignments: A discussion of methodologies. Working paper, Boston College, Boston.

Ball R, Brown P (1968) An empirical evaluation of accounting income numbers. *J. Accounting Res.* 6(2):159–178.

Barberis N, Shleifer A, Wurgler J (2005) Comovement. *J. Financial Econom.* 75(2):283–317.

Basu S (1997) The conservatism principle and the asymmetric timeliness of earnings. *J. Accounting Econom.* 24(1):3–37.

Banz RW (1981) The relationship between return and market value of common stocks. *J. Financial Econom.* 9(1):3–18.

Ben-David I, Franzoni F, Moussawi R (2018) Do ETFs increase volatility? *J. Finance* 73(6):2471–2535.

Bhattacharya A, O'Hara M (2018) Can ETFs increase market fragility? Effect of information linkages on ETF markets. Working paper, Cornell University, Ithaca, NY.

Boehmer B, Boehmer E (2003) Trading your neighbor's ETFs: Competition or fragmentation? *J. Banking Finance* 27(9):1667–1703.

Boehmer E, Kelley E (2009) Institutional investors and the informational efficiency of prices. *Rev. Financial Stud.* 22(9):3563–3594.

Boone A, White J (2015) The effect of institutional ownership on firm transparency and information production. *J. Financial Econom.* 117(3):508–553.

Boroujerdi R, Fogertey K (2015) ETFs: The rise of the machines. Report, Goldman Sachs Equity Research, New York.

Bushee B, Noe C (2000) Corporate disclosure practices, institutional investors, and stock return volatility. *J. Accounting Res.* 38(Suppl. 2000):171–202.

Chakravarty S (2001) Stealth-trading: Which traders' trades move stock prices? *J. Financial Econom.* 61(2):289–307.

Chan KC, Hamao Y, Lakonishok J (1991) Fundamentals and stock returns in Japan. *J. Finance* 46(5):1739–1764.

Chang Y-C, Hong H, Liskovich I (2015) Regression discontinuity and the price effects of stock market indexing. *Rev. Financial Stud.* 28(1):212–246.

Chen S, Huang Y, Li N, Shevlin T (2019) How does quasi-indexer ownership affect corporate tax-planning? *J. Accounting Econom.* 67(2):278–296.

Chordia T, Shivakumar L (2006) Earnings and price momentum. *J. Financial Econom.* 80(3):627–656.

Chordia T, Subrahmanyam A, Tong Q (2014) Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *J. Accounting Econom.* 58(1):41–58.

Collins DW, Kothari SP, Shanken J, Sloan RG (1994) Lack of timeliness and noise as explanations for the low contemporaneous return-earnings association. *J. Accounting Econom.* 18(3):289–324.

Cong L, Xu D (2016) Rise of factor investing: Asset prices, informational efficiency, and security design. Working paper, University of Chicago, Chicago.

Crane AD, Michenaud S, Weston JP (2016) The effect of institutional ownership on payout policy: A regression discontinuity design approach. *Rev. Financial Stud.* 29(6):1377–1408.

Da Z, Shive S (2016) Exchange traded funds and asset return correlations. *Eur. Financial Management* 24(1):136–168.

Diamond D, Verrecchia R (1991) Disclosure, liquidity, and the cost of capital. *J. Finance* 46(4):1325–1359.

Ernst & Young (2017) Reshaping around the investor: Global ETF research 2017. Accessed December 1, 2018, https://www.ey.com/Publication/vwLUAssets/ey-global-etf-survey-2017/$FILE/ey-global-etf-survey-2017.pdf.

Fama E, French K (2015) A five-factor asset pricing model. *J. Financial Econom.* 116(1):1–22.

Fama E, MacBeth J (1973) Risk, return, and equilibrium: Empirical tests. *J. Political Econom.* 81(3):607–636.

Ferreira M, Matos P (2008) The colors of investors' money: The role of institutional investors around the world. *J. Financial Econom.* 88(3):499–533.

Gillan SL, Starks LT (2003) Corporate governance, corporate ownership, and the role of institutional investors: A global perspective. *J. Appl. Finance* 13(2):4–22.

Glasserman P, Mamaysky H (2016) Market efficiency with micro and macro information. Working paper, Columbia University, New York.

Glosten L, Milgrom P (1985) Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *J. Financial Econom.* 14(1):71–100.

Gompers P, Metrick A (2001) Institutional investors and equity prices. *Quart. J. Econom.* 116(1):229–259.

Greenwich Associates (2016) Institutional investment in ETFs: Versatility fuels growth. Accessed January 1, 2016, https://www.greenwich.com/asset-management/institutional-investment-etfs-versatility-fuels-growth.

Greenwood R, Sosner N (2007) Trading patterns and excess comovement of stock returns. *Financial Anal. J.* 63(5):69–81.

Hamm SJW (2014) The effect of ETFs on stock liquidity. Working paper, Ohio State University, Columbus.

Harris L, Gurel E (1986) Price and volume effects associated with changes in the S&P 500: New evidence for the existence of price pressures. *J. Finance* 41(4):815–829.

Hasbrouck J (2003) Intraday price formation in U.S. equity index markets. *J. Finance* 58(6):2375–2399.

Hayn C (1995) The information content of losses. *J. Accounting Econom.* 20(2):123–153.

Healy PM, Hutton AP, Palepu KG (1999) Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Res.* 16(3):485–520.

Investment Company Institute (2018) *Investment Company Fact Book*, 58th ed. (Investment Company Institute, Washington, DC). https://www.ici.org/pdf/2018_factbook.pdf.

Israeli D, Lee C, Sridharan S (2017) Is there a dark side to exchange traded funds (ETFs)? An information perspective. *Rev. Accounting Stud.* 22(3):1048–1083.

Jiambalvo J, Rajgopal S, Venkatachalam M (2002) Institutional ownership and the extent to which stock prices reflect future earnings. *Contemporary Accounting Res.* 19(1):117–145.

Karmaziene E, Sokolovski V (2014) Exchange traded funds and the 2008 short-sale ban. Swedish House of Finance Research Paper, Swedish House of Finance, Stockholm.

Krause T, Ehsani S, Lien D (2014) Exchange traded funds, liquidity, and market volatility. *Appl. Financial Econom.* 24(24):1617–1630.

Kyle AS (1985) Continuous auctions and insider trading. *Econometrica* 53(6):1315–1335.

Lakonishok J, Shleifer A, Vishny R (1994) Contrarian investment, extrapolation, and risk. *J. Finance* 49(5):1541–1578.

Lambert R, Leuz C, Verrecchia R (2007) Accounting information, disclosure, and the cost of capital. *J. Accounting Res.* 45(2):385–420.

Lettau M, Madhavan A (2018) Exchange-traded funds 101 for economists. *J. Econom. Perspect.* 32(1):135–154.

Li F, Zhu Q (2016) Synthetic shorting with ETFs. Working paper, Hong Kong University of Science and Technology, Hong Kong.

Lundholm R, Myers L (2002) Bringing the future forward: The effect of disclosure on the returns-earnings relation. *J. Accounting Res.* 40(3):809–839.

Madhavan A (2016) *Exchange-Traded Funds and the New Dynamics of Investing* (Oxford University Press, Oxford, UK).

Madhavan A, Marchioni U, Li W, Du D (2014) Equity ETFs vs. index futures: A comparison for the fully-funded investor. *J. Index Investing* 5(2):66–75.

Nallareddy S, Ogneva M (2017) Accrual quality, skill, and the cross-section of mutual fund returns. *Rev. Accounting Stud.* 22(2):503–542.

Newey W, West K (1987) A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3):703–708.

Peng L, Xiong W (2006) Investor attention, overconfidence and category learning. *J. Financial Econom.* 80(3):563–602.

Piotroski J, Roulstone B (2004) The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *Accounting Rev.* 79(4):1119–1151.

Piwowar M (2017) SEC-NYU dialogue on exchange-traded products. Accessed December 1, 2018, https://www.sec.gov/news/speech/speech-piwowar-2017-09-08.

Ramaswamy S (2011) Market structures and systemic risks of exchange-traded funds. Working paper, Bank of International Settlements, Basel, Switzerland.

Shleifer A, Vishny R (1986) Large shareholders and corporate control. *J. Political Econom.* 94(3):461–488.

Shleifer A, Vishny R (1997) The limits of arbitrage. *J. Finance* 52(1):35–55.

Staer A (2014) Equivalent volume and comovement. Working paper, California State University, Fullerton.

Vijh A (1994) S&P 500 trading strategies and stock betas. *Rev. Financial Stud.* 7(1):215–251.

Wurgler J (2010) On the economic consequences of index-linked investing. Rosenfeld G, Lorsch J, Khurana R, eds. *Challenges to Business in the Twenty-First Century: The Way Forward* (American Academy of Arts and Sciences, Cambridge, MA).