Is there a Dark Side to Exchange Traded Funds (ETFs)? An Information Perspective

by

Doron Israeli, Charles M. C. Lee, and Suhas A. Sridharan**

First Draft: September 12, 2014

Current Draft: May 12, 2016

Abstract

We examine whether an increase in ETF ownership is accompanied by a decline in pricing efficiency for the underlying component securities. Our tests show an increase in ETF ownership is associated with: (1) higher trading costs (bid-ask spreads and price impact of trades); (2) an increase in "stock return synchronicity"; (3) a decline in "future earnings response coefficients"; and (4) a decline in the number of analysts covering the firm. Collectively, our findings support the view that increased ETF ownership can lead to higher trading costs and lower benefits from information acquisition. This combination results in less informative security prices for the underlying firms.

JEL Classifications: G11, G14, M41

Keywords: Exchange traded funds (ETFs); Uninformed and informed traders; Costly information; Trading costs; Market pricing efficiency; Informative prices.

** Israeli (israelid@idc.ac.il) is at the Arison School of Business, the Interdisciplinary Center (IDC), Herzliya, Israel; Lee (clee8@stanford.edu) is at the Graduate School of Business, Stanford University, and Sridharan (sridharan@emory.edu; 404-727-9431) is at the Goizueta Business School, Emory University (Atlanta, GA 30322). We gratefully acknowledge research assistance from Woo Young Park and Padmasini Venkatachari, as well as suggestions and comments from Ananth Madhavan, Ed Watts, and seminar participants at Emory University, Interdisciplinary Center (IDC) Herzliya, Tel Aviv University, UCLA, the University of Iowa, and Duke University.

I. Introduction

Traditional noisy rational expectations models with costly information feature agents who expend resources to become informed. These informed agents earn a return on their information acquisition efforts by trading with the uninformed, and as they do so, the information they possess is incorporated into prices.[1] In these models, the supply of uninformed traders adjusts to provide just sufficient reward for costly efforts in information acquisition and processing. The equilibrium between cost constraints faced by informed traders and gains from trading against the uninformed is reflected in the level of informational efficiency of security prices in the market. The inherent tension between the efficiency with which firm-specific information is being incorporated into stock prices, and the incentives needed to acquire that information and disseminate it, is central to understanding the informational content and role of security prices (e.g., Hayek 1945, Grossman 1989).

This paper employs exchange-traded fund (ETF) ownership data to examine the economic linkages between the market for firm-specific information, the market for individual securities, and the role of uninformed traders. Specifically, we study the influence of ETF ownership on the informational efficiency (or "price informativeness") of the individual component securities underlying the fund.[2] In frictionless markets, a firm's ownership structure should have little to do with the informational efficiency of its share price. However, as we argue below, market frictions related to information acquisition costs can cause ownership by ETFs to be a significant economic event, with direct consequences for the informational efficiency of the underlying securities.

Our central conjecture is that ETF ownership can influence a stock's informational efficiency through its impact on the supply of underlying securities available for trade, as well as the number of uninformed traders willing to trade these securities. As ETF ownership grows, an

[1] See for example, Grossman and Stiglitz (1980), Hellwig (1980), Diamond and Verrecchia (1981), Verrecchia (1982), Admati (1985), and Kyle (1985, 1989).

[2] We use the terms "price informativeness" or "informational efficiency" interchangeably. Both terms refer to the speed and efficiency with which price incorporates new information. Empirically, we use several proxies to measure informational efficiency, including "price synchronicity" (*SYNCH*), "future earnings response coefficients" (*FERC*), and the number of analysts covering a firm (*ANALYST*).

increasing proportion of the outstanding shares for the underlying security becomes "locked up" (held in trust) by the fund sponsor. Although these shares are available for trade as part of a basket transaction at the ETF-level, they are no longer available to traders who wish to transact on firm-specific information. Even more importantly, ETFs offer an attractive investment alternative for uninformed (or "noise") traders who would otherwise trade the underlying component securities.[3] As ETF ownership increases, uninformed traders in the underlying securities tend to migrate toward the ETF market. Over time, both effects create a steady siphoning of firm-level liquidity which in turn generates a disincentive for informed traders to expend resources to obtain firm-specific information.

We propose and test two hypotheses. First, we posit that as ETFs become larger holders of a firm's shares, *transaction costs for the underlying securities will increase*. This increase in trading costs is associated with a decrease in available liquidity for the component securities owned by ETFs. Second, we posit that these increased transaction costs will lead to *a general deterioration in the pricing efficiency of the underlying securities*. Specifically, we posit that the increased transaction costs will serve as a deterrent to traders who would otherwise expend resources on information acquisition about that stock. In other words, for firms that are widely-held by ETFs, the incentive for agents to seek out, acquire, and trade on firm-specific information will decrease. Over time, this will result in a general deterioration in the firm's information environment, and a reduction in the extent to which its stock price is able to quickly reflect firm-specific information.[4]

[3] A number of models predict noise investors will migrate to index-like instruments because their losses to informed traders are lower in these markets than in the market for individual securities (e.g., Rubinstein 1989; Subrahmanyam 1991; Gorton and Pennacchi 1993; Bhattacharya and O'Hara 2016; Cong and Xu 2016). Empirically, we have observed such a migration from actively managed assets to passively managed products, particularly ETFs. As of June 2015, total ETF trading already represented close to 28% of the total daily value traded on US equity exchanges (Pisani 2015).

[4] Note that the siphoning of liquidity from component securities can occur with other basket securities as well, such as open-end index funds. However, a key difference between ETFs and other index-linked open-end funds is that ETF shares are traded on organized exchanges throughout the day, while transactions with open-end funds occur only at the end of the day, and only at net asset value (NAV). Thus we expect ETFs to be more attractive instruments for noise (or speculative) traders, while index funds are better suited to longer term buy-and-hold investors. In section II, we explain in detail the implications of this difference for our tests.

To test these hypotheses, we conduct a series of tests using a cross-section of U.S. stocks between 2000 and 2014. Our research design makes use of panel-data based on firm-year observations.[5] To conduct these tests, we first collect end-of-year ETF ownership data for all firms. We then examine the effect of *changes* of in ETF ownership on the component securities': (1) trading costs and market liquidity, and (2) various proxies of firm-level pricing efficiency.

In our trading cost tests, we follow prior literature (Goyenko et al. 2009, Corwin and Schultz 2012, Amihud 2002) in using two proxies of firm liquidity – the relative bid-ask spreads, *HLSPREAD*, and an adjusted measure of price impact of trades, *ILLIQ_N*.[6] After controlling for firm size, book-to-market ratio, share turnover, return volatility, and overall level of institutional ownership, we find that an increase in ETF ownership is associated with an increase in average daily bid-ask spreads of the component securities, measured over the next year. In addition, we show an increase in ETF ownership is associated with lower market liquidity (higher price impact) in the underlying security over the next year.

Our tests show a one percentage point increase in ETF ownership is associated with an increase of 1.7% in the average bid-ask spreads over the next year. At the same time, a one percentage point increase in ETF ownership is associated with an increase of 4.2% in average absolute returns over the next year. These findings are consistent with Hamm (2014), who reports that increased ETF ownership is associated with an increase in the "Kyle Lambda" (a stock illiquidity measure) for the underlying component securities owned by these funds.[7]

To test the information-related effects of ETF ownership, we examine the effect of ETF ownership on two proxies for the extent to which stock prices reflect firm-specific information: (1) stock return synchronicity, *SYNCH* (the extent to which firm-specific stock return variation is

[5] We use annual holding periods to test our hypotheses because we expect the information-related effects of ETF ownership changes to be experienced gradually over time. Our inferences are the same if we use quarterly panels.
[6] For reasons detailed in section III, we decompose the Amihud (2002) measure of price impact of trades and investigate the effect of increased ETF ownership on the numerator of the Amihud (2002) measure, *ILLIQ_N*, controlling for the denominator of the Amihud (2002) measure, *ILLIQ_D*.
[7] Compared to Hamm (2014), we use alternative measures of stock liquidity, include different control variables, examine annual vs. quarterly observations, and use a more complete firm-level longitudinal data set. However, our main findings with respect to the effect of ETF ownership on stock liquidity are consistent with her results. Hamm (2014) does not examine the implications of ETF ownership on the informational efficiency of firm prices.

attributable to general market and related-industry movements), and (2) future earnings response coefficient, *FERC* (the association between current firm-specific returns and future firm-specific earnings). In addition, we examine whether an increase in ETF ownership is associated with a decline in the number of analysts covering the firm.[8]

Our results are broadly consistent with the information-related hypothesis. Specifically, we find that an increase in ETF ownership is accompanied by a decline in the pricing efficiency of the underlying component securities, as measured by either *SYNCH* or *FERC*. Our results indicate that a one-percentage point increase in ETF ownership is associated with 4% increase in return synchronicity. In addition, firms experiencing a one-percentage point increase in ETF ownership also experience a 14% reduction in the magnitude of their future earnings response coefficients. These results are robust to various model perturbations, as well as the inclusion of controls for institutional ownership and a host of other variables prescribed by prior literature (Roll 1988, Durnev et al. 2003, Piotroski and Roulstone 2004, Ettredge et al. 2005, Choi et al. 2011). We also find that an increase in ETF ownership is accompanied by a decline in the number of analysts covering the firm.

It is instructive to compare and contrast our results with the findings reported in a recent working paper by Glosten, Nallareddy, and Zou (2015; hereafter, GNZ). Like us, GNZ also examine the effect of ETF trading on the informational efficiency of underlying securities. However, they document an *increase* in information efficiency for firms with increased ETF trading. This evidence suggests that an increase in ETF ownership improves pricing efficiency in the underlying stock. At first blush, these findings seem at odds with ours. However, using the same filtering rules as GNZ, we are able to replicate their finding, and reconcile them with our own.

A key difference in the research design between the two studies is in the timing of the ETF trades. While we examine the effect of *past* changes in ETF ownership on *future* earning response coefficients (FERCs), GNZ's tests are focused on the effect of contemporaneous ETF

[8] These measures have been featured in prior literature on pricing efficiency (e.g., Roll 1988, Durnev et al. 2003, Piotroski and Roulstone 2004, Ettredge et al. 2005, Choi et al. 2011).

trading on current quarter earnings-response-coefficients (ERC). In other words, their study focuses the effect of *contemporaneous increases in ETF ownership* on the market's ability to incorporate *same-quarter* earnings. Whereas their study examines how current-quarter ETF ownership changes affect price discovery of current-quarter earnings, our study is focused on longer-term implications of ETF ownership changes for the informational environment of the underlying firms.

A number of studies in the market microstructure literature suggest that trading associated with the ETF-arbitrage mechanism can improve intraday price discovery for the underlying securities (Hasbrouck 2003, Yu 2005, Chen and Strother 2008, Fang and Sanger 2012, and Ivanov et al. 2013), particularly if the individual securities are less liquid than the ETF. The idea is that traders can respond to earnings news (particularly the macro-related component of earnings) more quickly by trading the lower cost ETF instrument. As a result, the price of the ETF may lead the price of the underlying securities in integrating this type of news. Hasbrouck (2003) provide some empirical evidence for this phenomenon using index futures. GNZ findings are consistent with this idea in that increases in ETF ownership in a given quarter are associated with higher *same-quarter* ERCs.

Applying the same data filters as GNZ, we also find a positive contemporaneous correlation between increases in ETF ownership and the market's ability to incorporate same-quarter earnings. However, we go a step further and show that this positive relation holds only when all three variables (stock returns, ETF changes, and earnings) are measured in the same quarter. As we lengthen the time lag between ETF changes and future earnings, the relation turns negative. Moreover, as we increase the time lag between past ETF changes and current returns, the negative relation becomes stronger. In other words, while same-quarter ETF trading seems to improve pricing efficiency, the more salient result over the longer run is that increases in ETF ownership lead to a deterioration in pricing efficiency for the underlying securities.

We also provide some interesting evidence on the difference between "macro-based" and "firm-specific" earnings. GNZ posit that increased ETF trading can enhance price discovery for "macro-based" earnings information. The idea is that, for this type of earnings news, informed traders would prefer to trade through the ETF, which is a low-cost venue. To test this conjecture, GNZ parse the earnings of each firm into a “macro-based” component and a “firm-specific” component. Their results show that increases in ETF ownership primarily improve the market’s ability to integrate “macro-based” earnings news. In their tests, the effect of ETF ownership changes on “firm-specific” ERCs is insignificant.

In contrast, our main hypothesis is that the cost of information arbitrage will increase with ETF ownership. While this effect should reduce firm-specific FERC, it could also reduce the macro-based FERC. This is because as ETF ownership increases, all investors (both informed and uninformed) face higher trading costs, and consequently have lower incentives to acquire and analyze information about the underlying securities. Therefore, over time, we would expect increased ETF ownership to be associated with lower FERCs on both the macro-based and firm-specific components of earnings.

Our results largely support this hypothesis. First, we replicate the GNZ result using quarterly and annual panels. Second, we show that, over time, the positive correlation between returns and same period earnings turns negative for both the “macro-based” and “firm-specific” component of earnings. In fact, we find that the negative impact of increased ETF ownership on firms’ FERC is generally more pronounced for the “firm-specific” component earnings. Taken together, our findings confirm the GNZ finding that ETF trading improves pricing discovery for same-quarter “macro-based” earnings. However, we also show this positive effect is short-lived. Over the longer term, the primary effect of increases in ETF ownership is to lower FERCs, particularly with respect to the firm-specific component of earnings.

These findings contribute to a growing literature on the economic consequences of basket or index-linked products. The rapid increase in index-linked products in recent years has attracted the attention of investors, regulators, and financial researchers.[9] A number of prior studies suggest that trading associated with the ETF-arbitrage mechanism can improve intraday price

[9] Sullivan and Xiong (2012) note that while passively managed funds represent only about one-third of all fund assets, their average annual growth rate since the early 1990’s is 26 percent, double that of actively managed assets. Much of the increase in passively managed assets has been in the form of ETFs. According to Madhavan and Sobczyk (2014) as of June, 2014 there were 5,217 global ETFs representing $2.63 trillion in total net assets. By June 2015, ETF trading was already 28% of the total daily value traded in US exchanges (Pisani 2015).

discovery for the underlying stocks (Hasbrouck 2003, Yu 2005, Chen and Strother 2008, Fang and Sanger 2012, and Ivanov et al. 2013). Other studies highlight concerns related to the pricing and trading of these instruments, including the more rapid transmission of liquidity shocks, higher return correlations among stocks held by same ETFs (Da and Shive 2013, Sullivan and Xiong 2012), greater systemic risk (Ramaswamy 2011), and elevated intraday return volatility both for the component stocks and for the entire market (Ben-David et al. 2014, Broman 2013, Krause et al. 2013), particularly in times of market stress (Wurgler 2010).

Our study adds a longer-term informational perspective to this debate. Adopting key insights from information economics (Rubinstein 1989, Subrahmanyam 1991, Gorton and Pennacchi 1993, Bhattacharaya and O'Hara 2016, Cong and Xu 2016), we present empirical evidence on how incentives in the market for information can affect pricing in the market for the underlying securities. Our results suggest that ETF ownership can lead to increased trading costs for market participants, which has further consequences for the amount of firm-specific information that is incorporated into stock prices. While the benefits of ETFs to investors are well understood (Rubinstein 1989), far less is known about other (unintended) economic consequences they may bring to financial markets. Our findings help highlight a potentially negative consequence of the ETFs.

Evidence presented in this study also provides support for a long-standing prediction of the noisy rational expectations literature. A number of models in this literature (Grossman and Stiglitz 1980, Hellwig 1980, Admati 1985, Diamond and Verrecchia 1981, Verrecchia 1982, and Kyle 1985, 1989) predict that when information is costly to acquire and process, informational efficiency of security prices will vary with the supply of uninformed investors willing to trade these securities. Using the emergence of ETFs, we link the siphoning of firm-level liquidity and an increase in trading costs to a reduction in the incentives for information acquisition, and hence lower pricing efficiency.

Lee and So (2015) argue that the study of market efficiency involves the analysis of a joint equilibrium in which all markets need to be cleared simultaneously. Specifically, supply must equal demand in the market for information about the underlying security, as well as in the market for the security itself. Our findings provide support for this view of market efficiency, and bring into sharp relief the close relationship between the market for component securities and the market for information about these securities.

The remainder of our study is organized as follows. In the next section, we provide some institutional details on ETFs. In section III, we develop our main hypotheses and outline our research design. Section IV reports the empirical findings, and section V concludes.

II. Exchange-traded funds (ETFs)

In the United States, ETFs are registered under the Investment Company Act of 1940 and are classified as open-ended funds or as unit investment trusts (UITs). Like open-end index funds, in a typical ETF, the underlying basket of securities is defined with the objective of mimicking the performance of a broad market index. But ETFs differ in some important respects from traditional open-ended funds. For example, unlike open-ended funds, which can only be bought or sold at the end of the trading day for their net asset value (NAV), ETFs can be traded throughout the day much like a closed-end fund.[10] In addition, ETFs do not sell shares directly to investors. Instead, they only issue shares in large blocks called "creation units" to authorized participants ("APs") who effectively act as market-makers.

Only the ETF manager and designated APs participate in the primary market for the creation/redemption of ETF shares. At the inception of the ETF, APs buy an appropriate basket of the predefined securities and deliver them to the ETF manager, in exchange for a number of ETF "creation units". Investors can then buy or sell individual shares of the ETF from APs in the secondary market on an exchange. Shares of the ETF trade during the day in the secondary market at prices that can deviate from their net asset value (NAV), but the difference is kept in line through an arbitrage mechanism in the primary market. For example, when an ETF is

[10] Specifically, unlike ETFs, open-ended do not provide a ready intraday market for deposits and redemptions with a continuous series of available transaction prices. Hence, investors may not know with sufficient certainty the cash-out value of redemption before they must commit it.

trading at a premium to an AP's estimate of value, the AP may choose to deliver the creation basket of securities in exchange for ETF shares, which in turn it could elect to sell or keep.

Notice that the creation/redemption mechanism in the ETF structure allows the number of shares outstanding in an ETF to expand or contract based on demand from investors. As Madhavan and Sobczyk (2014) observe, this creation/redemption mechanism means that "liquidity can be accessed through primary market transactions in the underlying assets, beyond the visible secondary market." This additional element of liquidity means that trading costs of ETFs are determined by the *lower bound* of execution costs in either the secondary or primary markets, a factor especially important for large investors." (p.3). In other words, unlike open-end funds, APs interested in accessing the assets represented by the ETF can now choose to trade either in the secondary ETF market (buy/sell the ETF shares directly), or in the primary market (buy/sell the basket securities).

For other (non-AP) investors, ETFs offer the convenience of a stock (ETFs can be bought and sold throughout the day, like common stocks) along with the diversification of a mutual fund or index funds (they give investors a convenient way to purchase a broad basket in a single transaction). Unlike open-end index funds (or other basket securities), ETFs do not require investors to deal directly with the fund itself. The most popular ETFs also tend to be much more liquid than the underlying securities, making them useful instruments for speculators and active traders.[11] Finally, adding to their appeal to active traders, ETF shares can also be borrowed and sold short.

In sum, ETFs possess many of the characteristics of what Rubinstein (1989) calls an "ideal market basket vehicle." In particular, ETFs (1) have a continuous market through time of basket sales and purchases (i.e., provide reliable cash-out prices prior to commitment to trade), (2) have low creation costs (i.e., trade execution costs incurred in the original purchase of components of

[11] It should be noted that ETFs are most likely to be successful when the underlying securities are relatively less liquid or difficult to borrow (thus creating an equilibrium demand for the ETF shares, with its lower trading costs). For example, the highly popular small-cap ETF, IWM, is based on the Russell 2000 index. While the underlying securities are typically less liquid (i.e. they represent the 2,000 stocks in the Russell Index that are below the largest 1,000), IWM itself is over $26 billion in size and trades at extremely low costs.

the underlying basket and organization costs), (3) enhance tax benefits obtained from positions in the individual components of the basket (there is no taxation of unrealized profits; unlike open-ended mutual funds, which typically fund shareholder redemptions by selling portfolio securities, ETFs usually redeem investors in-kind), (4) are offered in small enough units to appeal to small investors (not just to large institutional investors), and (5) remove all basket-motivated trading away from the individual securities or risks comprising the basket. These characteristics make ETFs especially attractive to active noise (uninformed) traders who would otherwise trade the underlying securities. Accordingly, we conjecture that uninformed traders will gravitate towards ETFs and away from the underlying stocks, with attendant consequences for the trading costs and pricing efficiency of the underlying securities.

III. Hypothesis development and research design

The primary goal of this study is to investigate whether an increase in the proportion of firm shares held by ETFs is associated with a decline in the pricing efficiency of the underlying component securities. To address this question we identify two central dimensions of a firm's information environment: (1) transactions costs of market participants, and (2) the extent to which stock prices reflect firm-specific information. We then make predictions about the effects of ETF ownership on each of these dimensions and construct tests to evaluate these predictions.

We first posit that ETFs serve as attractive substitutes to the underlying component securities for uninformed traders. Because of the trading benefits offered by ETFs, especially to uninformed investors, we expect uninformed investors to gravitate towards ETFs and away from the underlying stocks (Milgrom and Stokey 1982, Rubinstein 1989). As uninformed traders shift towards trading ETFs and away from trading the underlying securities, transactions costs for trading the underlying component securities will increase (Subrahmanyam 1991, Gorton and Pennacchi 1993, Mahavan and Sobczyk 2014). The increase in transactions costs will deter market participants from engaging in firm-specific information gathering activities and will lead to less informative stock prices in the firm-specific component (Grossman and Stiglitz 1980; Admati 1985). Based on the reasoning outlined above, we raise the following hypotheses:

H1: An increase in ETF ownership is associated with higher trading costs for the underlying component securities.

H2: An increase in ETF ownership is associated with deterioration in the pricing efficiency of the underlying component securities.

To test H1, we examine the relation between change in ETF ownership and changes in two proxies of liquidity that capture trading costs: (1) bid-ask spreads, and (2) price impact of trades (Goyenko et al. 2009). To investigate to the relation between ETF ownership and bid-ask spreads, we estimate the following regression:[12]

$$\Delta HLSPREAD_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \sum_k \beta_k \, \Delta Controls_{it-1} + \sum_j \beta_j \, INDST\_FE_i + \sum_l \beta_l \, YEAR\_FE_t + \epsilon_{it} \qquad (1)$$

In Eq. (1), the Δ operator indicates a change in the value of a particular variable. For example, $\Delta HLSPREAD_{it}$ is the difference between firm *i*'s measure of $HLSPREAD$ during year *t* and its value in year *t-1*. $HLSPREAD_{it}$, is the Corwin and Schultz (2012) annual high-low measure of bid-ask spread for firm *i* over year *t*. We use this measure of bid-ask spread as a proxy for trading costs because it is much less time and data-intensive to calculate than intraday bid-ask spread measures, and because Corwin and Schultz (2012) demonstrate that it outperforms the Roll (1984), Lesmond et al. (1999), and Holden (2009) techniques for measuring bid-ask spreads. However, none of the main results or inferences change when these other bid-ask spread measures are used.

[12] We test our hypotheses using annual panels because we expect the effect of increased ETF ownership to manifest itself gradually over time after an increase in ETF ownership. Figure 2 presents a sample construction timeline for the key empirical variables used in our tests. Most of our analyses are done using *annual* changes in ETF ownership, returns, and earnings (Panel A). However, in our replication and reconciliation of the GNZ results, we used *quarterly* data (Panel B) to match their analyses.

The variable of interest in Eq. (1), ΔETF_{it-1}, is the change in the percentage of firm *i*'s shares held by all ETFs from the end of year *t-2* to the end of year *t-1*. Our first hypothesis predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus*, an increase in ETF ownership is associated with an increase in bid-ask spread. Change in ETF ownership may be correlated with overall change in institutional ownership and prior research suggests there might be a relation between institutional ownership and bid-ask spreads.[13] To isolate the effect of change in ETF ownership on stock liquidity and to ensure that our results are not confounded by the relation of ETF ownership with institutional ownership, we include $\Delta INST_{it-1}$ directly in Eq. (1) as an additional control variable.[14] $\Delta INST_{it-1}$ is the change in the percentage of firm *i*'s shares held by all institutions from the end of year *t-2* to the end of year *t-1*.

In Eq. (1), $Controls_{it-1}$ represents a vector of firm and industry related control variables nominated by prior literature. We control for the change in the log of market value of equity [Δ*LN(MVE)*] during year *t-1* because larger firms generally have smaller bid-ask spreads. Prior studies also find that bid-ask spreads increase with the return volatility and decrease with the share turnover (Copeland and Galai 1983). Accordingly, we control for the change in the annualized standard deviation of daily returns during year *t-1* [Δ*STD(RET)*] and the change in average share turnover from year *t-2* to year *t-1* (Δ*TURN*). We include the change in book to market ratio (Δ*BTM*) during year *t-1* as a control for financial distress and/or growth opportunities (Fama and French 1992, Lakonishok et al. 1994). Finally, to control for time and industry trends in bid-ask spreads, we include year and industry fixed effects.[15] Our first hypothesis (H1) predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus,* increases in ETF ownership are associated with subsequent increases in bid-ask spreads.

As an additional test of H1, we also examine the price impact of trades as an alternative measure of firms' market liquidity or transaction costs. The Amihud (2002) illiquidity ratio, *ILLIQ*, defined as the ratio of average daily absolute returns to average daily dollar volume, is a

[13] Prior research on the relation between bid-ask spreads and institutional ownership is mixed. Glosten and Harris (1988) suggest that higher levels of concentrated institutional ownership will increase bid-ask spreads, while higher levels of dispersed institutional ownership might encourage competition that reduces bid-ask spreads.
[14] Our inferences are the same when we use the residual from the regression model $\Delta ETF_{it} = \beta_0 + \beta_1 \Delta INST_{it} + \varepsilon_{it}$ as a measure of change in ETF ownership that is orthogonal to the change in the level of institutional ownership.
[15] The industry fixed effects are defined based on the 48 Fama and French (1997) industry classification.

well-accepted proxy for price impact of trades (Goyenko et al. 2009). However, in the context of this paper, using *ILLIQ* as originally defined complicates our analyses. This is because prior literature (Hasbrouck 2003, Yu 2005, Chen and Strother 2008, Fang and Sanger 2012, and Ivanov et al. 2013) shows that ETF ownership can affect both the numerator of the illiquidity ratio (the average daily absolute returns) and the denominator of the illiquidity ratio (the average daily dollar volume).

To mitigate this problem, we decompose *ILLIQ* into two components (the numerator and the denominator) and estimate the following regression:

$$\Delta ILLIQ\_N_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \beta_3 \Delta ILLIQ\_D_{it} + \textstyle\sum_k \beta_k \, \Delta Controls_{it-1} + \sum_j \beta_j \, INDST\_FE_i + \sum_l \beta_l \, YEAR\_FE_t + \epsilon_{it} \tag{2}$$

$ILLIQ\_N_{it}$ is the daily absolute return for firm *i* averaged over all the trading days in year *t*. The dependent variable in Eq. (2), $\Delta ILLIQ\_N_{it}$, is the change in $ILLIQ\_N_{it}$ from year *t-1* to year *t*. $ILLIQ\_D_{it}$ is the daily dollar volume for firm *i* averaged over all the trading days in year *t* and $\Delta ILLIQ\_D_{it}$ is the change in $ILLIQ\_D_{it}$ from year *t-1* to year *t*. ΔETF_{it-1} and $\Delta INST_{it-1}$ are as defined above. $\Delta Controls_{it-1}$ denotes several control variables measured as of the end of year *t-1*. Specifically, it includes the log of market value of equity [*LN(MVE)*] as of the end of year *t-1*, because we expect larger firms to exhibit smaller price impact of trades. In addition, $\Delta Controls_{it-1}$ contains the change in book-to market-ratio (*ΔBTM*) during year *t-1* to control for the effects of financial distress and/or growth opportunities. In our estimation of Eq. (2) we also include year and industry fixed effects. Our hypothesis predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus* increases in ETF ownership are associated with increases in the price impact of trades (and hence lower liquidity or higher trading costs for market participants).

H2 states that an increase in ETF ownership is associated with deterioration in pricing efficiency of the underlying component security. We test this hypothesis using two proxies for the extent to which stock prices reflect firm-specific information: (1) stock return synchronicity, *SYNCH*, and (2) future earnings response coefficient, *FERC*.

SYNCH is a measure of the extent to which firm-level return variation is explained by general and related-industry return variation. Roll (1988) posits that when greater levels of firm-specific information are being impounded into stock prices, the magnitude of the stock return synchronicity measure decreases. Wurgler (2000), Durnev et al. (2003), Durnev et al. (2004), and Piotroski and Roulstone (2004) use this insight and provide evidence in support of it in a variety of settings. Because stock return synchronicity is negatively related to the amount of firm-specific information embedded in stock price, based on H2, we predict that changes in ETF ownership lead to positive changes in stock return synchronicity.

To estimate firm-specific measures of stock return synchronicity, $SYNCH_{it}$, we follow the methodology outlined by Durnev et al. (2003). First, for each firm-year observation we obtain the adjusted coefficient of determination (adjusted R^2) by regressing daily stock returns on the current and prior day's value-weighted market return (*MKTRET*) and the current and prior day's value-weighted Fama and French 48 industry return (*INDRET*):

$$\begin{aligned} RET_{id} = \alpha + \beta_1 MKTRET_d + \ &\beta_2 MKTRET_{d-1} \\ + \beta_3 INDRET_d + \beta_4 INDRET_{d-1} &+ \epsilon_{id} \end{aligned} \tag{3}$$

In Eq. (3), RET_{id} is firm *i's* stock return on day *d*, $MKTRET_d$ is the value-weighted market return on day *d*, and $INDRET_d$ is the value-weighed return of firm *i*'s industry, defined using the Fama-French 48 classifications, on day *d.*[16] Eq. (3) is estimated separately for each firm-year, using daily returns for firm *i* over the trading days in year *t*, with a minimum of 150 daily observations.

Next, for each firm-year observation we calculate the annual measure of stock return synchronicity, $SYNCH_{it}$, as the logarithmic transformation of R_{it}^2 to create an unbounded

[16] We adopt this model of returns to measure firm-specific adjusted R^2 (and, consequently, synchronicity) because it is the most frequently used in the literature (e.g., Piotroski and Roulstone 2004, Hutton et al 2010, Chan and Chan 2014). To ensure that our inferences are not affected by the method chosen to estimate firm specific adjusted R^2 we also estimate synchronicity using the measures outlined in Crawford et al (2012) and Li et al. (2014). Our inferences are unchanged by these alternate measurement techniques.

continuous measure of synchronicity (Piotroski and Roulstone 2004, Hutton et al. 2010, Crawford et al. 2012, Hutton et al. 2010):[17] $SYNCH_{it} = log\left(\frac{R_{it}^2}{1-R_{it}^2}\right)$. High values of the $SYNCH_{it}$ measure indicate that a greater fraction of firm-level return variation is explained by general market and related-industry return variation.

To test whether an increase in ETF ownership is accompanied by a decline in the amount of firm-specific information that is being impounded into stock prices we estimate the following equation: [18]

$$\Delta SYNCH_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \sum_k \beta_k \Delta Controls_{it-1} + \sum_j \beta_j INDST\_FE_i + \sum_l \beta_l YEAR\_FE_t + \epsilon_{it} \quad (4)$$

In Eq. (4), $\Delta SYNCH_{it}$ is the difference between firm *i*'s measure of $SYNCH$ during year *t* and its value in year *t-1*. $\Delta Controls_{it-1}$ indicates several annual change measures that prior research suggests are associated with changes in stock return synchronicity. In particular, following Jin and Myers (2006), we control for changes in the skewness of firm *i*'s returns over year *t-1* (*ΔSKEW*). In addition, since Li et al. (2014) show that synchronicity is often confounded with systematic risk, we include the annual change in CAPM beta as a control for a firm's systematic risk. As additional controls, we include annual changes during year *t-1* in the log of market value of equity [*ΔLN(MVE)*], book-to-market ratio *(ΔBTM)*, average share turnover (*ΔTURN)*, and year and industry fixed effects. ΔETF_{it-1} and $\Delta INST_{it-1}$ are defined and measured as in Eqs. (1) and (2). Our second hypothesis predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus,* increases in ETF ownership are associated with increases in stock return synchronicity.

Our second proxy for the extent to which stock prices reflect firm specific information is

[17] In computing $SYNCH_{it}$ we exclusively use adjusted R_{it}^2 values. Following Crawford et al. (2012), we truncate the sample of adjusted R_{it}^2 values at 0.0001.
[18] Consistent with prior literature, we examine whether an increase in ETF ownership is accompanied by a decline in the extent to which firm-level stock returns reflect future firm-specific earnings, using both levels and changes specifications. This is mainly because, as opposed to other equations, the dependent variable in FERC tests [i.e., Eqs. (5), (5a), and (5b)] is firm-level stock returns, which is stationary over time.

the future earnings response coefficient, which measures the extent to which current stock returns reflect firm-specific future earnings. To test whether an increase in ETF ownership is accompanied by a decline in the extent to which firm-level stock returns reflect future firm-specific earnings, we follow prior literature (e.g., Kothari and Sloan 1992, Collins et al. 1994, Choi et al. 2011) and estimate several versions of the following regression model:

$$RET_{it} = \beta_1 EARN_{it-1} + \beta_2 EARN_{it} + \beta_3 EARN_{it+1} + \beta_4 ETF_{it-1} + \beta_5 INST_{it-1} \quad (5)$$
$$+ \beta_6 ETF_{it-1} \times EARN_{it-1} + \beta_7 ETF_{it-1} \times EARN_{it} + \beta_8 ETF_{it-1} \times EARN_{it+1}$$
$$+ \sum_k \beta_k\, Controls_{it} + \sum_j \beta_j\, INDST\_FE_i + \sum_l \beta_l\, YEAR\_FE_t + \epsilon_{it}$$

In Eq. (5), RET_{it} represents firm-level stock returns during year *t*, and $EARN_{it-1}$, $EARN_{it}$, and $EARN_{it+1}$ denote firm-level net income before extraordinary items during years *t-1*, *t*, and *t+1*, scaled by market value of equity. The coefficient β_3 measures the relation between current firm-level stock returns and future firm-level earnings; prior research refers to this coefficient as the "future earnings response coefficient" (*FERC*) and offers it as a measure of the extent to which current stock returns reflect/predict future earnings (e.g., Ettredge et al. 2005, Choi et al. 2011).

To address our research question we include as explanatory variables the level of ETF ownership (ETF_{it-1}) at the end of year *t - 1* as well as the interaction between the level of ETF ownership and past, current, and future earnings ($ETF_{it-1} \times EARN_{it\pm j}$). Our hypothesis predicts that the coefficient on the interaction of ETF ownership with current and future earnings is negative, indicating that *FERCs* are lower for firms with higher ETF ownership (i.e., β_7 and, more importantly, β_8 are negative).

As in previous equations, $Controls_{it}$ denotes a number of control variables as suggested by prior research. Specifically, following Collins et al. (1994), we control for future firm-level stock returns, RET_{it+1}, to address the potential measurement error induced by using actual future earnings as a proxy for expected future earnings. In addition, to account for the effect of a firm's growth on the ability of its stock returns to reflect future earnings, we control for total assets growth from year *t-1* to year *t*, $ATGROWTH_t$. Also, we control for the fact that firms experiencing losses are expected to have lower *FERCs* by including an indicator variable, $LOSS_t$, that equals one if the firm experiences a loss in year t+1 (i.e., $EARN_{it+1} < 0$) and 0 otherwise. $Controls_{it}$ also includes the natural logarithm of market value of equity at the end of year *t*. Our second hypothesis predicts that the coefficient β_7 and, more importantly, β_8 are negative.

We also examine how FERCs vary with changes in ETF ownership by decomposing the level of ETF ownership at the end of period *t-1* into the sum of the level of ETF ownership at the end of period *t-2* and the change in ETF ownership during period *t-1*:

$$ETF_{t-1} = ETF_{t-2} + \Delta ETF_{t-1}$$

Thus, we re-estimate Eq. (5) using ETF_{t-2} and ΔETF_{t-1} in lieu of ETF_{t-1}:

$$\begin{aligned} RET_{it} = {} & \beta_1 EARN_{it-1} + \beta_2 EARN_{it} + \beta_3 EARN_{it+1} + \beta_4 ETF_{it-2} + \\ & \beta_5 \Delta ETF_{it-1} + \beta_6 ETF_{it-2} \times EARN_{it-1} + \beta_7 ETF_{it-2} \times EARN_{it} + \beta_8 ETF_{it-2} \times \\ & EARN_{it+1} + \beta_9 \Delta ETF_{it-1} \times EARN_{it-1} + \beta_{10} \Delta ETF_{it-1} \times EARN_{it} + \\ & \beta_{11} \Delta ETF_{it-1} \times EARN_{it+1} + \beta_{12} INST_{it-1} + \textstyle\sum_k \beta_k \, Controls_{it} + \\ & \textstyle\sum_j \beta_j \, INDST\_FE_i + \sum_l \beta_l \, YEAR\_FE_t + \epsilon_{it} \end{aligned} \tag{5a}$$

In estimating Eq. (5a), we expect that β_7 and, more importantly, β_8 – the coefficients on the interactions of lagged levels of ETF ownership with current and future earnings (i.e., $ETF_{t-2} \times EARN_t$ and, more importantly, $ETF_{t-2} \times EARN_{t+1}$) – as well as β_{10} and, more importantly, β_{11} – the coefficients on the interactions of lagged changes in ETF ownership with current and future earnings (i.e., $\Delta ETF_{t-1} \times EARN_t$ and, more importantly, $\Delta ETF_{t-1} \times EARN_{t+1}$) − are negative.

The impact of ETF ownership changes on pricing efficiency may differ for "macro-based" (systematic or aggregate) and "firm-specific" (or idiosyncratic) components of earnings. To test this conjecture, we follow the procedure in GNZ and decompose total earnings into "macro-based" and "firm-specific" components by estimating the following regression:

$$EARN_{it} = \beta_1 EARNMKT_t + \beta_2 EARNIND_{it} + \epsilon_{it} \quad (6)$$

In Eq. (6), $EARNMKT_t$ is the size-weighted average of year t earnings before extraordinary items for all firms with available earnings information in Compustat. $EARNIND_t$ is the size-weighted average of year t earnings before extraordinary items for all firms with the same Fama-French 48 industry classification.

For each firm-year, we define the systematic or aggregate portion of earnings ($EARNAGG_{it}$) as the fitted value from the annual estimation of Eq. (6). The residual portion is defined as the idiosyncratic or firm-specific portion of earnings ($EARNFIRM_{it}$). Using these components of earnings, we estimate the following modified version of Eq. (5):

$$\begin{aligned}
RET_{it} = {} & \beta_1 EARNAGG_{it-1} + \beta_2 EARNFIRM_{it-1} + \beta_3 EARNAGG_{it} + \beta_4 EARNFIRM_{it} \\
& + \beta_5 EARNAGG_{it+1} + \beta_6 EARNFIRM_{it+1} + \beta_7 ETF_{it-2} + \beta_8 \Delta ETF_{it-1} \qquad (5b) \\
& + \beta_9 ETF_{it-2} \times EARNAGG_{it-1} + \beta_{10} ETF_{it-2} \times EARNAGG_{it} + \beta_{11} ETF_{it-2} \times EARNAGG_{it+1} \\
& + \beta_{12} ETF_{it-2} \times EARNFIRM_{it-1} + \beta_{13} ETF_{it-2} \times EARNFIRM_{it} + \beta_{14} ETF_{it-2} \times EARNFIRM_{it+1} \\
& + \beta_{15} \Delta ETF_{it-1} \times EARNAGG_{it-1} + \beta_{16} \Delta ETF_{it-1} \times EARNAGG_{it} + \beta_{17} \Delta ETF_{it-1} \times EARNAGG_{it+1} \\
& + \beta_{18} \Delta ETF_{it-1} \times EARNFIRM_{it-1} + \beta_{19} \Delta ETF_{it-1} \times EARNFIRM_{it} + \beta_{20} \Delta ETF_{it-1} \times EARNFIRM_{it+1} \\
& + \beta_{21} INST_{it-1} + \textstyle\sum_k \beta_k Controls_{it} + \sum_j \beta_j INDST_FE_i + \sum_l \beta_l YEAR_FE_t + \epsilon_{it}
\end{aligned}$$

With the exception of $EARNAGG_{it}$ and $EARNFIRM_{it}$, all variables in Eq. (5b) remain as defined in Eq. (5a). Our second hypothesis predicts the coefficients on the interaction of lagged ETF ownership measures with current and future firm-specific earnings (β_{13}, $\beta_{14,}$ β_{19} and β_{20}) will be negative. In addition, if the deterioration in pricing efficiency is more pronounced for firm-specific earnings, these coefficients will be more negative than the analogous coefficients on interactions involving $EARNAGG_{it}$ (β_{10}, $\beta_{11,}$ β_{16} and β_{17}).

As an additional test of H2, we examine how ETF ownership relates to the number of analysts covering the firm during a year. Our hypothesis predicts higher ETF ownership will lead to lower incentives for information acquisition for the underlying securities. To the extent that analysts are drawn to firms that are more attractive to active (i.e., informed) investors, we expect that firms with increases in ETF ownership will be associated with reductions in analyst coverage. To test this, we estimate several versions of the following equation:

$$\Delta ANALYST_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \sum_k \beta_k \, \Delta Controls_{it-1} + \sum_j \beta_j \, INDST\_FE_i + \sum_l \beta_l \, YEAR\_FE_t + \epsilon_{it} \tag{7}$$

In Eq. (7), $\Delta ANALYST_{it}$ is the change from year *t-1* to year *t* in the number of unique analysts on I/B/E/S providing forecasts of firm *i*'s one-year-ahead earnings. As before, $\Delta Controls_{it-1}$ represents annual changes in a number of control variables, measured as the change in the level of each variable from year *t-2* to year *t*-1, which are suggested by prior literature. Barth et al. (2001) demonstrate that firms with large research and development expenses or intangible assets experience greater analyst coverage. To control for this effect, we include the annual change in the proportion of research and development expenses relative to total operating expenses (ΔRD_F_{it-1}) and the annual change in the proportion of intangible assets relative to total assets ($\Delta INTAN\_F_{it-1}$) as controls. Following Lang and Lundholm (1996), we also control for annual change in return volatility [$\Delta STD(RET)_{it-1}$]. To capture the effect of stock return momentum on levels of analyst coverage, we include prior firm-level 6-month equity returns (MOM_{it-1}), measured as of the end of year *t*-1. Eq. (7) also includes controls for firm size [*ΔLN(MVE)*], change in book-to-market ratio (*ΔBTM*), and change in share turnover (*ΔTURN*). Our second hypothesis predicts that the coefficient β_1 is negative, indicating that, *ceteris paribus,* changes in ETF ownership are associated with decline in number of analysts covering a firm.

To control for potential time-series as well as cross-sectional correlations between firm-specific measures, we base our inferences from all equations on *t*-statistics calculated using standard errors clustered by both firm and year (e.g., Gow et al. 2010). All variables used in the estimation of Eqs. (1) to (7) are also defined in Appendix A.

IV. Empirical Analyses

IV.1 Sample construction and descriptive statistics

We determine year-end ETF ownership by first using CRSP, Compustat, and OptionMetrics data bases to identify all ETFs traded on the major U.S. exchanges. Specifically, we identify ETFs as securities on CRSP with a share code of '73' and securities on Compustat or OptionMetrics with a issue type of '%'. After identifying candidate ETFs, we obtain for each ETF the reported equity holdings from the Thomson Financial S12 database. For some ETFs, the Thomson Financial S12 database does not provide regular reporting of equity holdings. In these instances, we hand collect additional holdings data from Bloomberg Financial. ETFs without any reported holding data in the Thomson Financial database or Bloomberg Financial are excluded from the sample. This process yields a sample of 443 unique ETFs. Appendix B provides a list of the 10 largest ETFs in our sample, ranked based on the average assets under management.

Using the annual panel of holdings for each ETF we define, for every stock in a given year, the ETF ownership variable (*ETF*) as the aggregate number of shares held by all ETFs divided by total number of shares outstanding in that year. We repeat this process for every firm-year between 2000 and 2014 to construct our panel. Our sample begins in 2000 because it is the first year with sufficient variation in ETF ownership to conduct our analyses. Our sample ends in 2014 due to data availability constraints. All firm-years with no reported ETF ownership in the sample period are included in the sample with $ETF_{it} = 0$.

Figure 1 reports the average ETF ownership across firms for each year of our sample. The figure reveals a significant increase in average ETF ownership over our sample period, from roughly 1% in 2000 to nearly 5.5% in 2014. Perhaps even more telling is the rapid increase in the dollar value of ETF trading as a percentage of total exchange value traded. During June 2015, total ETF trading represented close to 28% of the total daily exchange value traded (Pisani 2015). This represents a 35% increase from June 2014. Clearly ETFs have quickly become an important vehicle for traders in the equity market.

We obtain market-related data on all US-listed firms from CRSP and accounting data from Compustat. To be included in our sample, each firm-year observation must have information on stock price, number of shares outstanding, and book value of equity. We also require sufficient data to calculate the standard deviation of daily returns and average share turnover within each firm year. We restrict our analyses to firms with non-negative book-to-market ratios in every year of our sample period. This results in a sample of 39,863 firm-years and 7,489 unique firms. In some of our analyses, we also require data on annual analyst coverage. In analyst coverage analyses, our sample size is reduced to 29,562 firm-year observations. The number of observations included in each regression varies according to data availability.

Panel A of table 1 presents descriptive statistics for the main variables used in the analyses. Of particular interest for our analyses is the level of ETF holding, measured as a percentage of total shares outstanding held by all ETFs. The mean (median) percentage ETF ownership is 3.31% (2.52%). This is much lower than the level of institutional ownership, which has a mean (median) of 57.78% (62.50%). We also observe consistently larger annual changes in institutional ownership relative changes in ETF ownership. The mean (median) change in ETF ownership is 48 (27.7) basis points while the mean (median) change in institutional ownership is 264 (109) basis points. The distributional statistics of both ETF and institutional ownership in our sample are consistent with prior literature (Hamm 2014, Jiambalvo 2002). Nevertheless, we expect the two measures to differ in many important respects and have different effects on measures of trading costs and pricing efficiency.

Panel A also reveals that Δ*ILLIQ_N* and Δ*ILLIQ_D*, the two components of changes in the Amihud (2002) illiquidity ratio, have notably different variances. Δ*ILLIQ_N* is very narrowly distributed with a standard deviation of .801, while Δ*ILLIQ_D* exhibits a significantly larger standard deviation of 15.661. This difference provides further support for our decision to decompose the Amihud (2002) ratio into its two components in an attempt to estimate the effect of changes in ETF ownership on Δ*ILLIQ_N* controlling for Δ*ILLIQ_D*.

Table 1, panels B and C present Pearson and Spearman correlation coefficients between the key levels and changes of variables in our empirical analysis. In our sample, Δ*ETF* is positively correlated with changes in book to market ratio (Pearson coef. = 0.022) and turnover (Pearson coef. = 0.07). Panel B reveals that Δ*ETF* is positively correlated with our two proxies of trading costs, *HLSPREAD* (Pearson coef. = .171) and Δ*ILLIQ_N* (Pearson coef. = .193). Consistent with our hypotheses, Δ*ETF* is also positively correlated with Δ*SYNCH* (Pearson coef. = .080) and negatively correlated with Δ*ANALYST* (Pearson coef. = -.005).

IV.2 Testing H1: ETF ownership and trading costs of market participants

Tables 2 and 3 present regression summary statistics from the estimation of Eqs. (1) and (2) which are designed to test our first hypothesis using two measures of liquidity that capture trading costs and various model specifications.

Column 1 of table 2 reveals that change in bid-ask spread, Δ*HLSPREAD*, exhibits the expected relations with our control variables. Δ*HLSPREAD* is negatively associated with increases in firm size (coef. = -0.044, *t*-stat. = -1.67), positive associated with increases in the BTM ratio (coef. = 0.051, *t*-stat. = 2.42), and positively associated with increases in return volatility (coef. = 0.001, *t*-stat = 1.86). According to column 1 of table 2, changes in ETF ownership are positively related to changes in bid-ask spreads (coef. = 0.016, *t*-stat. = -2.41). This finding supports our hypothesis that, ceteris paribus, the trading costs of market participants increase with changes in ETF ownership. To ensure that the results in column 1 are not confounded by the relation between changes in institutional ownership with Δ*HLSPREAD* and with Δ*ETF*, we estimate column 1, controlling for *ΔINST*. Column 2 reveals that consistent with H1, there exists a significantly positive association between changes in *ETF* ownership and changes in *HLSPREAD* (coef. = 0.017, *t*-stat. = 2.51). The coefficient on institutional ownership itself is slightly negative and insignificant (coef. = -0.000, *t*-stat. = -0.58).

Results from estimating Eq. (2) are presented in table 3 and provide additional evidence on the association between measures of ETF ownership and trading costs. Column 1 reveals that Δ*ILLIQ_N* has a strong positive association with changes in book to market ratio (coef. = 0.172, *t*-stat. = 4.20) and changes in *ILLIQ_D* (coef. = 0.004, *t*-stat. = 5.38). The results in column 1 also reveal a strong positive relationship between changes in ETF ownership and changes in *ILLIQ_N* (coef. = 0.042, *t*-stat. = 2.55), suggesting that increases in ETF ownership are associated with increases in the Amihud illiquidity measure. Column 2 of table 3 reveals that controlling for *INST* does not alter this observed relationship. Taken together, the results presented in tables 2 and 3 provide strong evidence in support for H1 that an increase in ETF ownership is accompanied by an increase in trading costs for market participants.

IV.3 Testing H2: ETF ownership and the deterioration of pricing efficiency

IV.3a Synchronicity and FERC tests

Tables 4 and 5 present regression summary statistics from the estimation of Eqs. (4) and (5), which are designed to test our second hypothesis using two proxies for the extent to which stock returns reflect firm-specific information. Table 4 presents the summary statistics from the estimation of two versions of Eq. (4), which models the relation between changes in ETF ownership and changes in annual stock return synchronicity. Column 1 reveals that the changes in synchronicity, Δ*SYNCH*, exhibit the expected relations with our control variables. Consistent with prior research (Li et al. 2014), Δ*SYNCH* is positively associated with increases in firm size (coef. = 0.595, *t*-stat. = 8.78). Δ*SYNCH* has a negative association with changes in systematic risk, Δ*BETA* (coef. = -.704, *t*-stat. = -14.85). Columns 1 and 2 reveals that changes in ETF ownership are significantly positively related to stock return synchronicity (coef. = .090, *t*-stat. = 3.70 and 3.67), even after controlling for changes in institutional ownership. This supports our hypothesis that increases in ETF ownership are associated with a deterioration of pricing efficiency for the underlying component securities.

Table 5 presents regression summary statistics from the estimation of Eqs. (5), (5a), and (5b) which is designed to examine the relation between ETF ownership and the extent to which current firm-level returns reflect future firm-specific earnings. In all of our FERC tests, our measure of ETF ownership is defined from 0 to 1 (rather than from 0 to 100) to be more similar in magnitude to our earnings and returns measures. Consistent with prior literature, we observe a positive future earnings response coefficient in both columns (coef. = 0.015, *t*-stat. = 4.80 and

4.32). Consistent with H2, column 1 of panel A reveals that the interactions of current and future earnings with ETF ownership carry significantly negative coefficients (coef. = -3.662, *t*-stat. = -2.60 and coef. = -0.212, *t*-stat. = -4.64). This suggests that controlling for $INST_{t-1}$ and a host of other variables prescribed by prior literature [e.g., $LOSS_t$, $ATGROWTH_t$, RET_{t+1}, $LN(MVE)_{t-1}$] firms with higher levels of ETF ownership experience lower future earnings response coefficients. In other words, firm-level returns of firms with higher levels of ETF ownership incorporate less future earnings-related information. Column 2 of panel A presents summary statistics from the estimation of Eq. (5a) in which the level of ETF ownership is split into lagged level and most recent period change. The results in column 2 provide further support for H2 by showing that the interaction of current and future earnings with changes in ETF ownership are also significantly negative (coef. = -3.636, *t*-stat. = -2.68 and coef. = -0.195, *t*-stat. = -2.07), suggesting that magnitude of future earnings response coefficients shrink as ETF ownership increases.

Taken together, the results presented in tables 4 and 5 indicate that an increase in ETF ownership is associated with increase in the co-movement of firm-level stock returns with general market and related-industry stock returns, and with a decline in the predictive power of current firm-level stock returns for future earnings. These two findings support our second hypothesis that stock prices of firms with high ETF ownership are impounding less firm-specific information.

IV.3b Alternative earnings response tests

In a contemporaneous study, Glosten, Nallareddy, and Zou (2015; hereafter GNZ) explore the impact of ETF trading activity on the response of returns to contemporaneous earnings news. Using same-quarter changes in ETF ownership as a proxy for ETF trading activity, they demonstrate that ETF trading is associated with a stronger association of returns to contemporaneous earnings. They further document that the effect is concentrated in the association of returns to the systematic component of earnings news. Their overall conclusion from these findings is that ETF trading improves informational efficiency.

Given the difference in the conclusions of the GNZ study relative to our own, we take several steps to reconcile their findings with ours. The first key difference between their setting and ours is rooted in the time lag between measurement of ETF ownership changes, returns, and earnings. Because GNZ's focus is on understanding the impact of contemporaneous ETF trading, they measure changes in ETF ownership and returns over the exact same quarter. In contrast, we are interested in understanding the long-run implications of ETF ownership on pricing efficiency, so we measure levels and changes of ETF ownership *prior* to the start of the returns measurement window. To verify that this shift in ETF measurement window is a key driver of the differences between our results and those of GNZ, we estimate Eq. (5b) while measuring changes in ETF ownership and earnings contemporaneously with returns, as GNZ does.

Column 1 of table 5, panel B presents the results of this estimation. These results show that the coefficients on the variables of interest (the interaction of ΔETF with contemporaneous aggregate and firm earnings) are positive, consistent with the findings reported by GNZ. To ensure that the effect documented by GNZ does not subsume those reported in our FERC analyses, we also estimate Eq. (5b) including both the GNZ measurement of ΔETF and our original measurement of ΔETF. The summary statistics from this estimation are presented in column 2 of table 5, panel B. This test shows that all our prior findings on a reduction in FERCs with increased ETF ownership continue to hold, after controlling for the GNZ variables. Specifically, all eight of the interaction terms between ETF ownership and future earnings have negative coefficients. In particular, the coefficient of lagged changes in ETF ownership with future firm-specific earnings is significantly negative (coef. = -0.172, *t*-stat = -3.26), as is the interaction of the lagged level of ETF ownership with future firm-specific earnings (coef. = -0.136, *t*-stat = -2.02). These results support our hypothesis that increases in ETF ownership lead to slower incorporation of firm-specific earnings information into stock prices.

The analyses presented in table 5, panel B show the importance of lagging the changes in ETF ownership variable, which is a key difference between our study and GNZ. However, there are other research design differences that might have contributed to the differences in results (such as annual vs. quarterly data, the measurement window for future earnings, the choice of control variables, and their use of seasonally adjusted earnings). To address these differences more completely, we re-estimate the main results from GNZ's Eq. (3) following their research design and sample construction.

The summary statistics from this estimation are presented in column I of table 6. Their main variable of interest is the interaction of the contemporaneous quarterly change in ETF ownership (δETF_t) with current period seasonally adjusted earnings ($SEARN_t$). For firm *i* in quarter *t*, $SEARN_{it}$, seasonally adjusted earnings news, is defined as the difference between quarter *t* and quarter *t-4* earnings before extraordinary items, scaled by stock price at the beginning of quarter *t*. Consistent with the results reported by GNZ, our estimate of the coefficient on the interaction $\delta ETF_t \times SEARN_{it}$ is positive and significantly different from zero (coef = 0.438, *t* stat.= 2.458).

Having successfully replicated the results reported by GNZ, we further explore how research design modifications affect the gap between their results and ours. The estimates in columns II through VII of table 6 are from re-estimations of GNZ's Eq. (3) with modifications to the measurement windows of δETF and *SEARN*. Specifically, in column II we hold all other aspects of the research design constant (and consistent with GNZ), but allow for the measurement of δETF to take place prior to the return measurement window rather than in the same quarter as the stock returns. In other words, we shift from δETF_t to δETF_{t-1}. Making this shift causes the magnitude of the coefficient on the interaction of δETF_{t-1} with *SEARN* to decline (coef. = 0.386 versus 0.438) and also causes the statistical significance to fall (*t*-stat = 1.729 versus 2.458).

The next modification we examine is the effect of measuring annual changes in ETF ownership rather than quarterly. We accomplish this by summing quarterly changes over the prior four quarters. We use this four-quarter sum of ETF ownership changes ($\delta ETF_{SUM\ t-1\ to\ t-4}$) instead of the contemporaneous ETF ownership change in the estimation results presented in column III of table 6. The results in column III reveal that measuring ETF changes annually notably changes the inferences of the test, as the interaction of $\delta ETF_{SUM\ t-1\ to\ t-4}$ with *SEARN* is positive (coef = .101) but not significantly different from zero (*t* stat = .767).

In columns IV through VII we maintain use of the four-quarter sum of quarterly ETF ownership changes ($\delta ETF_{SUM\ t-1\ to\ t-4}$) as an annualized measure of change in ETF ownership. The variation in columns IV through VII arises from shifting forward the measurement of *SEARN*. In column IV (V, VI), *SEARN* is defined as the quarterly earnings for quarter $t+1$ ($t+2$, $t+3$) relative to the returns measurement window. In each iteration of the estimation, the coefficient on the interaction of $\delta ETF_{SUM\ t\text{-}1\ to\ t\text{-}4}$ with *SEARN* grows increasingly negative. For example, Column VI shows that when examining three-quarter-ahead earnings, the interaction of $\delta ETF_{SUM\ t\text{-}1\ to\ t\text{-}4}$ with $SEARN_{t+3}$ bears a significantly negative coefficient (coef. = -.568, t-stat= -1.809). This negative interaction effect is consistent with our second hypothesis.

In column VII we sum the quarterly earnings for quarters *t* through $t+3$ to approximate an annual version of future earnings ($SEARN_{SUM\ t\ to\ t+3}$). Column VII of Table 6 presents estimations where the combined modifications to the measurement of ETF ownership changes and earnings bring the overall specification close to our main FERC test, Eq. (5). Specifically, we use $SEARN_{SUM\ t\ to\ t+3}$ as an approximation of future annual earnings and $\delta ETF_{SUM\ t\text{-}1\ to\ t\text{-}4}$ as an approximation of annual change in ETF ownership *prior* to the returns measurement window. The results in column VII further support our second hypothesis; the coefficient on the interaction of $\delta ETF_{SUM\ t\text{-}1\ to\ t\text{-}4}$ with $SEARN_{SUM\ t\ to\ t+3}$ is -0.198 (t-stat = -1.848). This indicates that with larger increases in ETF ownership over the past year, current returns capture less information about earnings over the next year.

IV.3c Analyst coverage tests

Finally, Table 7 presents regression summary statistics from the estimation of Eq. (7), designed to examine the effect of ETF ownership changes on analyst coverage. The control variables largely have the predicted signs. For example, Column 1 shows that changes in analyst coverage, $\Delta ANALYST$, are higher among firms with larger increases in research and development expenses (coef. = 1.965, *t*-stat. = 2.79), and firms with larger increases in intangible assets (coef. = 1.386, *t*-stat. = 3.77). These results confirm prior studies that found demand for firm-specific information is higher among such firms.

In columns I and II the coefficients on ΔETF are positive but not significantly different from zero (coef. = 0.042 and 0.017, *t*-stat. = 1.24 and 0.46). These results suggest that increases in ETF ownership in over year t-1 have no significant effect on analyst coverage. However, the coefficient on $\Delta INST$ presented in column II is positive and highly significant (coef. = 0.013, *t*-stat. = 9.06), suggesting that an increase in institutional ownership over year t-1 does tend to increase analyst coverage in year t. These are consistent with our intuition as well as prior research on the relation between analyst coverage and institutional ownership.

To explore the possibility that ETF ownership might have a more long-run effect on analyst coverage, we include the level of ETF ownership as an additional explanatory variable in the estimation results presented in columns III and IV. The results show that the year t-2 level of ETF ownership exhibits a strong negative relationship with changes in analyst coverage in year t. For example, column IV shows that higher ETF ownership is associated with lower analyst coverage (coef. = -0.025, t-stat. = -3.25), after controlling for both the level and the change in lagged institutional ownership.

In sum, over shorter windows (i.e. with a one-year lag) we find no reliable evidence linking changes in ETF ownership to the number of analysts covering a firm. Over longer horizons, our results do suggest that increased ETF ownership is associated with lower analyst coverage. However, we hasten to point out that as the time gap between the measurement of ETF ownership and analyst coverage increases, so does the likelihood that some other confounding factor is at work. Thus we think our analyst coverage results should be interpreted cautiously. At best, our findings are consistent with the idea that analysts slowly respond to changes in the information environment associated with from changes in ETF ownership.

V. Summary and concluding remarks

In this study, we use changes in ETF ownership to examine the economic linkages between the market for firm-specific information, the market for individual securities, and the role of uninformed traders. The market for ETFs has grown dramatically in the past decade and ETFs now constitute an attractive investment alternative, especially for those interested in diversified strategies. By focusing on the natural growth of exchange-traded funds over the past decade, we study how changes in the composition of a firm's investor base impacts its share pricing efficiency.

Prior theoretical work offers two predictions on the possible impact of ETF ownership. First, a number of studies in the market microstructure literature suggest that trading associated with the ETF-arbitrage mechanism can improve intraday price discovery for the underlying stocks (Hasbrouck 2003, Yu 2005, Chen and Strother 2008, Fang and Sanger 2012, and Ivanov et al. 2013). This line of inquiry suggests that increased ETF ownership can lead to improved pricing efficiency in the underlying securities.

On the other hand, the noisy rational expectations literature suggests a possible negative relation between ETF ownership and pricing efficiency. A number of models in this literature (Grossman and Stiglitz 1980, Hellwig 1980, Admati 1985, Diamond and Verrecchia 1981, Verrecchia 1982, and Kyle 1985, 1989) predict that with costly information, pricing efficiency will be a function of the total supply of uninformed investors willing to trade the security. To the extent that uninformed investors migrate away from underlying component securities as ETF ownership increases, these models predict that increased ETF ownership can lead to a decline in pricing efficiency.

Our study examines, and provides support for, both predictions. We first demonstrate that an increase in ETF ownership is associated with an increase in firms' trading costs. This is consistent with the idea of uninformed traders exiting the market of the underlying security in favor of the ETF. As uninformed traders exit and trading costs rise, we posit that pricing efficiency will decline. Consistent with this prediction, we find that increases in ETF ownership are associated with increases in stock return synchronicity and decreases in future earnings response coefficients (FERCs).

Consistent with the microstructure literature and Glosten, Nallareddy, and Zou (2015), we find increased ETF activity is leading to an improvement in same-quarter ERC. More importantly, we show that as the time lag between current period returns and future earnings is increased, this positive coefficient turns gradually negative. This negative relationship becomes stronger as we increase the time lag between the current period returns and prior period ETF changes. In other words, after allowing one to four quarters for changes in ETF ownership to take hold, FERCs will be significantly lower as a function of the increase in past period ETF ownership. These results are largely consistent with predictions from the noisy rational expectations literature.

In more detailed analyses, we also provide some evidence on the difference between "macro-based" and "firm-specific" earnings. GNZ find a positive relation between ETF changes and pricing efficiency, and show the effect is primarily driven by better price integration of "macro-based" earnings. In contrast, we find that the negative impact of increased ETF ownership on FERC is more pronounced for "firm-specific" earnings. Taken together, these findings suggest that ETF trading may improve pricing discovery for same-quarter "macro-based" earnings, but that over the longer term (beginning with the next quarter), increases in ETF ownership actually lead to lower FERCs, particularly with respect to firm-specific earnings news.

Our findings contribute to a growing literature on the economic consequences of basket or index-linked products. The rapid increase in index-linked products in recent years has attracted the attention of investors, regulators, and financial researchers. Our study adds a longer-term informational perspective to this debate. Adopting key insights from information economics, we present empirical evidence on how incentives in the market for information can affect pricing in the market for the underlying securities. Our results suggest that ETF ownership can lead to increased trading costs for market participants, with further consequences for the amount of firm-specific information that is incorporated into stock prices.

Lee and So (2015) argue that the study of market efficiency involves the analysis of a joint equilibrium in which all markets need to be cleared simultaneously. Specifically, supply must equal demand in the market for information about the underlying security, as well as in the market for the security itself. In the same spirit, Pedersen (2015) argues that financial markets should be viewed as "efficiently inefficient" – that is, neither perfectly efficient nor completely inefficient. Rather, Pedersen notes, markets are "inefficient enough that money managers can be compensated for their costs through the profits of their trading strategies, and efficient enough that the profits after costs do not encourage additional active investing." (Pedersen, jacket cover).

Our study provides support for this view of market efficiency, and our main findings bring into sharp relief the close relationship between the market for component securities and the market for information about these securities.

References

Agarwal, P. 2007. Institutional Ownership and Stock Liquidity. Working paper.

Admati, A. R. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53: 629-657.

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

Barth, M.E., R. Kasznik, and M. McNichols. 2001. Analyst Coverage and Intangible Assets. *Journal of Accounting Research 39*: 1-34.

Ben-David, I., F. Franzoni, and R. Moussawi. 2014. Do ETFs Increase Volatility? Working Paper. Ohio State University, University of Lugano and the Swiss Finance Institute, and University of Pennsylvania, January.

Broman, M. S. 2013. Excess Co-movement and Limits-to-Arbitrage: Evidence from Exchange-Traded Funds. Working paper, York University.

Bhattacharya, A., and M. O'Hara. 2016. Can ETFs increase market fragility? Effect of information linkages on ETF markets. Working paper. Cornell University, March 10.

Chan, K. and Y.C. Chan. 2014. SEO Underpricing, Stock Price Synchronicity and Analyst Coverage. *Journal of Financial Economics*, forthcoming.

Chen, G. and T. S. Strother. 2008. On the contribution of index exchange traded funds to price discovery in the presence of price limits without short selling. Working paper. University of Otago, New Zealand. February.

Choi, J., L. A. Myers, Y. Zang, and D. A. Ziebart, 2011. Do management EPS forecasts allow returns to reflect future earnings? Implications for the continuation of management's quarterly earnings guidance. *Review of Accounting Studies* 16: 143-182.

Collins, D., S. Kothari, J. Shanken, and R. Sloan. 1994. Lack of Timeliness versus Noise as Explanations for Low Contemporaneous Return-Earnings Association. *Journal of Accounting & Economics* 18: 289-324.

Cong, L. W., and D. Xu. 2016. Rise of ETFs: Informational efficiency and smart betas. Working paper. University of Chicago, April.

Copeland, T and D. Galai. 1983. Information Effects on the Bid-Ask Spread. *The Journal of Finance* 38: 1457-1469.

Corwin, S. and P. Schultz. 2012. A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *The Journal of Finance* 67: 719-759.

Crawford, S. S., D. T. Roulstone, and E. C. So, 2012. Analyst Initiations of Coverage and Stock Return Synchronicity. *The Accounting Review* 87: 1527-1553.

Da, Z. and Shive, S. 2013. Exchange-Traded Funds and Equity Return Correlations. Working paper, University of Notre Dame.

Diamond, D. W., and R. E. Verrecchia, 1981, Information aggregation in noisy rational expectations model. *Journal of Financial Economics* 9: 221-235.

Durnev, A., R. Morck, and B. Yeung. 2004. Value-enhancing capital budgeting and firm-specific stock return variation. *The Journal of Finance* 59: 65–105.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin. 2003. Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research* 41: 797–836.

Ettredge, M., S.Y. Kwon, and D. Smith. 2005. The Impact of SFAS no. 131 Business Segment Data on the Market's Ability to Anticipate Future Earnings. *The Accounting Review* 80 (3): 773 – 804.

Fang, Y., and G. C. Sang. 2012. Index price discovery in the cash market. Working paper. Louisiana State University, January 16th.

Glosten, L. and L. Harris. 1999. Estimating the components of the bid/ask spread. *Journal of Financial Economics*: 123-142.

Glosten, L., S. Nallareddy, and Y. Zou. 2015. ETF trading and informational efficiency of underlying securities. Working paper. Columbia University. October.

Gorton, G. B., and G. G. Pennacchi, 1993. Security Baskets and Index-Linked Securities. *The Journal of Business* 66: 1-27.

Gow, I., G. Ormazabal, and D. Taylor. 2010. Correcting for Cross-Sectional and Time-Series Dependence in Accounting Research. *The Accounting Review* 85: 483 – 512.

Goyenko, R. Y., G. W. Holden, and C. A. Trzcinka, 2009. Do liquidity measures measure liquidity? *Journal of Financial Economics* 92, 153-181.

Grossman S. 1989. *The informational role of prices*, MIT Press: Cambridge, MA.

Grossman, S., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70: 393-408.

Hamm, S. J. W., 2014. The Effects of ETFs on Stock Liquidity. Working paper. Ohio State University.

Hasbrouck, J., 2003. Intraday price formation in U.S. equity index markets. *The Journal of Finance* 58: 2375-2399.

Hayek, F. A., 1945. The Use of Knowledge in Society. *The American Economic Review* 35: 519-530.

Hellwig, M. R., 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22, 477-498.

Holden, C., 2009. New low-frequency spread measures. *Journal of Financial Markets* 12: 778-813.

Ivanov, S. I., F. J. Jones, and J. K. Zaima. 2013. Analysis of DJIA, S&P 500, S&P 400, NASDAQ 100 and Russell 2000 ETFs and their influence on price discovery. *Global Finance Journal* 24, 171-187.

Jiambalvo, J., S. Rajgopal, and M. Venkatachalam. 2002. Institutional Ownership and the Extent to which Stock Prices Reflect Future Earnings. *Contemporary Accounting Research* 19: 117-145.

Jin, L., and S. Myers. 2006. R-squared around the world: New theory and new tests. *Journal of Financial Economics* 79: 257–292.

Kothari, S. and R. Sloan. 1992. Information in Prices About Future Earnings: Implications for Earnings Response Coefficients. *Journal of Accounting and Economics* 15: 143-171.

Kyle, A. S., 1985. Continuous auctions and insider trading. *Econometrica* 53, 1315-1335.

___________, 1989. Informed speculation with imperfect competition. *Review of Economic Studies* 56, 317-356.

Krause, T., S. Ehsani, and D. Lien. 2013. Exchange traded funds, liquidity, and market volatility. Working paper. University of Texas at San Antonio. January 11.

Lang, M. and R. Lundholm. 1996. Corporate Disclosure Policy and Analyst Behavior. *The Accounting Review* 71: 467–92.

Li, B., S. Rajgopal, and M. Venkatachalam. 2014. R^2 and idiosyncratic risk are not interchangeable. *The Accounting Review,* forthcoming.

Lee, C. M. C., and E. So. 2015. Alphanomics: The Informational Underpinnings of Market Efficiency. *Foundations and Trends in Accounting* 9 (No. 2-3): 59-258.

Lesmond, D., J. Ogden, and C. Trzcinka. 1999. A new estimate of transactions costs. *Review of Financial Studies* 12: 1113–1141.

Li, B., S. Rajgopal, and M. Venkatachalam. 2014. R2 and Idiosyncratic Risk are not Interchangeable. *The Accounting Review*, forthcoming.

Madhavan, A. and A. Sobczyk. 2014. Price dynamics and liquidity of exchange-traded funds. Working paper. BlackRock Inc., August 27.

Milgrom, P. and N. Stokey. 1982. Information, trade and common knowledge. *Journal of Economic Theory* 26: 17-27.

Pedersen, L. H. 2015. *Efficiently Inefficient: How Smart Money Invests & Market Prices Are Determined.* Princeton University Press: Princeton NJ.

Piotroski, J. D. and D. T. Roulstone, 2004. The influence of analysis, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *The Accounting Review* 79: 1119–1151.

Pisani, R., 2015. Here is why ETFs are a growing part of total trading value, CNBC, July 2. http://www.cnbc.com/2015/07/02/heres-why-etfs-are-a-growing-part-of-total-trading-value.html

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance* 39: 127–1139.

Rubinstein, M., 1989. Market Basket Alternatives. *Financial Analysts Journal* 45: 20-29.

Stambaugh, R. F., 2014. Presidential Address: Investment Noise and Trends. *The Journal of Finance* 69: 1415-1453.

Subrahmanyam, A., 1991. A Theory of Trading in Stock Index Futures. *The Review of Financial Studies* 4: 17-51.

Verrecchia, R., 1982. Information acquisition in a noisy rational expectations economy. *Econometrica* 50, 1415-1430.

Wurgler, J., 2000. Financial markets and the allocation of capital. *Journal of Financial Economics* 58: 187–214.

Appendix A: Variable Definitions

Variable:		Definition
$ANALYST_{it}$	=	The number of unique analysts in I/B/E/S that provide forecasts of year *t* earnings for firm *i*
$ATGROWTH_{it}$	=	Growth in assets for firm *i* from year *t-1* to year *t*.
$BETA_{it}$	=	The β_1 coefficient from the firm-year estimation of the model $$(RET - R_f)_{id} = \alpha + \beta_1\left(MKT - R_f\right)_d + \epsilon_{id}$$ where $(RET - R_f)_{id}$ is firm *i's* return less the risk-free rate on day *d* and $(MKT\text{-}R_f)_d$ is the value-weighted market return less the risk-free rate on day *d*. The model is estimated using daily returns over the trading days in the year *t*, with a minimum of 150 trading days.
BTM_{it}	=	The book to market ratio of firm *i* at the end of year *t*
$EARN_{it}$	=	Earnings before extraordinary items for firm *i* in year *t* scaled by market value of equity
$EARNAGG_{it}$	=	The fitted value from the firm-year estimation of the model $$EARN_{it} = \beta_1 EARNMKT_t + \beta_2 EARNIND_{it} + \epsilon_{it}$$ where $EARNMKT_t$ is the size-weighted average of year *t* earnings before extraordinary items for all firms with available earnings information in Compustat and $EARNIND_t$ is the size-weighted average of year *t* earnings before extraordinary items for all firms with the same Fama-French 48 industry classification.
$EARNFIRM_{it}$	=	The residual value from the firm-year estimation of the model $$EARN_{it} = \beta_1 EARNMKT_t + \beta_2 EARNIND_{it} + \epsilon_{it}$$ where $EARNMKT_t$ is the size-weighted average of year *t* earnings before extraordinary items for all firms with available earnings information in Compustat and $EARNIND_t$ is the size-weighted average of year *t* earnings before extraordinary items for all firms with the same Fama-French 48 industry classification.
ETF_{it}	=	The percentage of firm *i's* common shares outstanding held by ETFs at the end of year *t*
$HLSPREAD_{it}$	=	The Corwin and Schultz (2012) measure of bid-ask spread for firm *i* in year *t*
$ILLIQ\_N_{it}$	=	The average over year *t* of absolute daily equity returns for firm *i* (the numerator of the Amihud (2002) illiquidity ratio)
$ILLIQ\_D_{it}$	=	The average over year *t* of dollar volume for firm *i* (the denominator of the Amihud (2002) illiquidity ratio)
$INST_{it}$	=	The percentage of firm *i's* common shares outstanding held by institutions at the end of year *t*

$INTAN\_F_{it}$	=	The ratio of intangible assets to total assets for firm *i* in year *t*
$LN(MVE)_{it}$	=	The natural logarithm of firm *i*'s market value of equity at the end of year *t*
$LOSS_{it}$	=	Indicator variable equaling one if firm *i* experienced a loss (defined as $EARN < 0$) in year *t*
MOM_{it}	=	Cumulative firm *i* stock returns for months -12 to -6 relative to the year *t* end date.
MTB_{it}	=	The ratio of market value of equity to book value of equity for firm *i* at the end of quarter *t*
RD_F_{it}	=	The ratio of research and development expenses to total operating expenses for firm *i* in year *t*
RET_{it}	=	The annual return for firm *i* in year *t*
$SEARN_{it}$	=	Seasonally adjusted quarterly earnings, defined as $$\frac{EARN_{it} - EARN_{it-4}}{P_{it-1}}$$ where $EARN_{it}$ is earnings before extraordinary items for firm *i* in quarter *t* and P_{t-1} is firm *i*'s stock price at the end of quarter *t-1*
$SKEW_{it}$	=	The skewness of firm *i*'s daily returns over year *t*
$SLOSS_{it}$	=	Indicator variable equaling one if $SEARN_{it} < 0$ in quarter *t*
STD_{it}	=	Standard deviation of firm *i*'s earnings per share excluding extraordinary items over the 20 quarters prior to quarter *t*
$STD(RET)_{it}$	=	The standard deviation of firm *i*'s daily returns over year *t*
$SYNCH_{it}$	=	A logarithmic transformation of R^2_{it} defined as $log\left(\frac{R^2_{it}}{1-R^2_{it}}\right)$, R^2_{it} is estimated separately for each firm-year as described in section III.
$TURN_{it}$	=	The ratio of the average number of firm *i*'s shares traded in year *t* to firm *i*'s total common shares outstanding in year *t*
Δ	=	The annual change operator
δ	=	The quarterly change operator

Appendix B: Sample ETFs ranked by assets under management

This appendix provides a list of the 10 largest ETFs in our sample, ranked using the average assets under management (AUM). Each year we compute AUM as of the last trading day of the calendar year.

Rank	Ticker	Fund name	Average AUM (in millions)
1	SPY	SPDR S&P 500 Trust ETF	$86,971.45
2	IVV	iShares Core S&P 500 ETF	$21,915.20
3	VTI	Vanguard Total Stock Market ETF	$14,140.61
4	IWM	iShares Russell 2000 ETF	$12,361.84
5	IWF	iShares Russell 1000 Growth Index ETF	$10,312.08
6	IWD	iShares Russell 1000 Value Index ETF	$8,965.49
7	DIA	SPDR Dow Jones Industrial Average ETF	$8,428.39
8	IJH	iShares Core S&P Mid-Cap ETF	$7,200.01
9	VOO	Vanguard S&P 500 ETF	$6,838.45
10	MDY	Spdr S&P Midcap 400 ETF	$6,783.54

Figure 1: ETF ownership by year

This chart plots, by fiscal year, the average percentage of shares outstanding held by ETFs for firms in our sample. The horizontal axis indicates the year and the vertical axis indicates the magnitude of ETF ownership. Our methodology for calculating ETF ownership is outlined in Section IV of the paper.

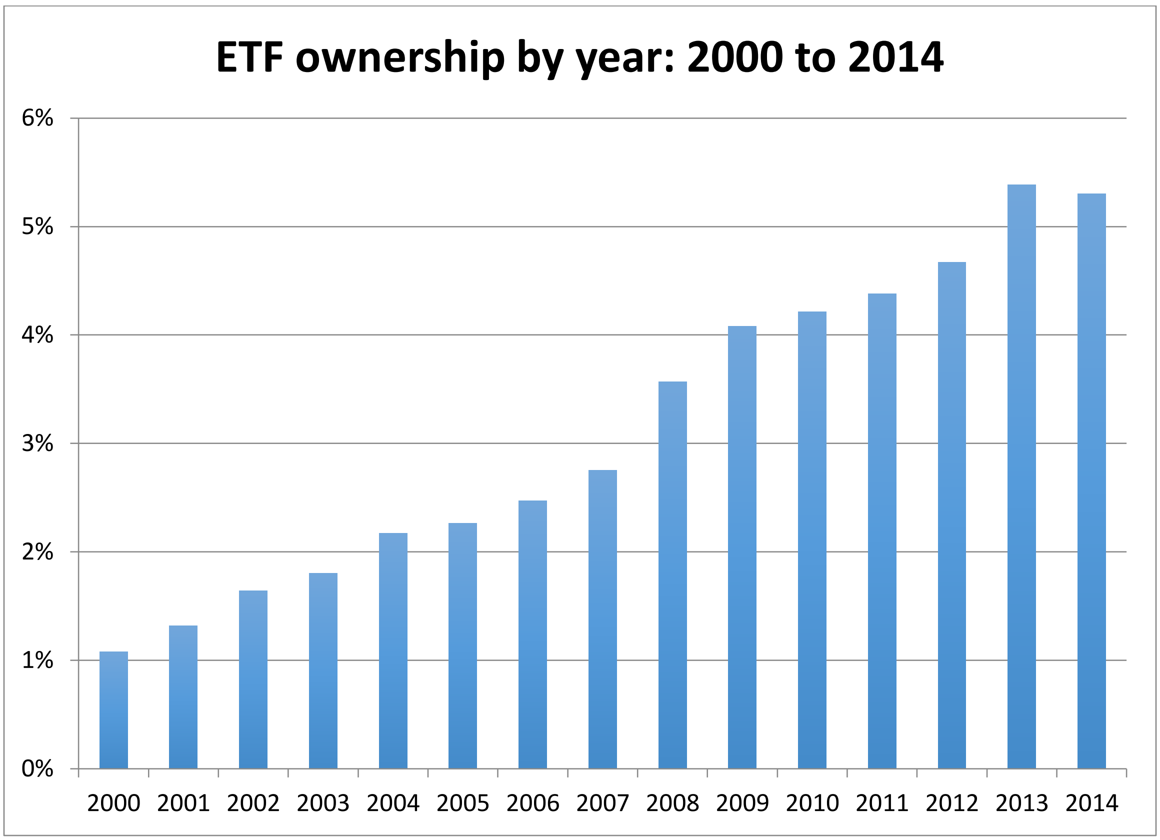
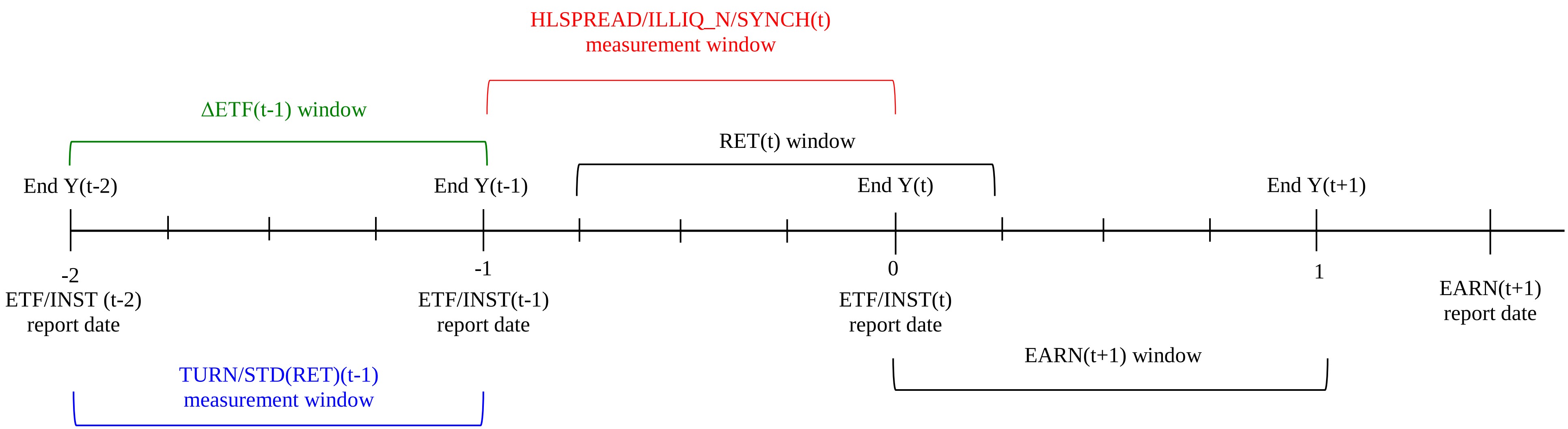


Figure 2: Sample construction timeline

Panel A: Annual sample



Panel B: Quarterly sample for GNZ replication

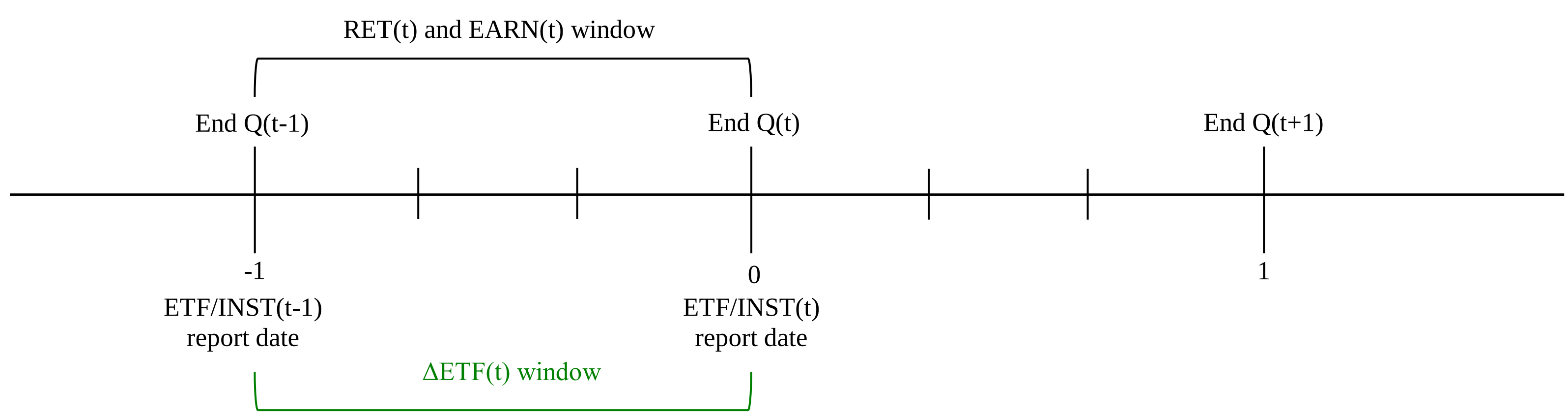


Table 1: Sample Description

Panel A: Univariate statistics

This panel presents univariate statistics for the key variables in our sample. Variable definitions are provided in Appendix A.

	N	MEAN	SD	Q1	Q2	Q3
ANALYST	29562	7.963	6.987	3	6	11
ATGROWTH	29970	22.835	59.931	-1.162	12.41	31.674
EARN	29970	-0.001	0.271	0.012	0.047	0.069
ETF	39863	3.306	2.972	1.223	2.52	4.707
INST	39863	57.775	31.256	30.833	62.497	85.46
LN(MVE)	39863	13.021	2.013	11.543	12.937	14.354
LOSS	29970	0.218	0.413	0	0	0
RET	29970	0.206	0.612	-0.14	0.113	0.401
ΔANALYST	29562	0.177	2.627	-1	0	1
ΔBETA	32536	0.036	0.418	-0.196	0.035	0.267
ΔBTM	39863	0.011	0.394	-0.117	0	0.123
ΔETF	39863	0.48	1.091	-0.001	0.277	0.879
ΔHLSPREAD	39863	-0.022	0.445	-0.231	-0.042	0.137
ΔILLIQ_D	39863	1.741	15.661	-0.326	0.018	1.454
ΔILLIQ_N	39863	-0.083	0.801	-0.501	-0.118	0.238
ΔINST	39863	2.64	11.579	-2.583	1.09	6.163
ΔLN(MVE)	39863	0.054	0.501	-0.189	0.08	0.327
ΔSKEW	32536	-0.026	1.423	-0.627	-0.032	0.568
ΔSTD(RET)	39863	-1.794	30.486	-15.074	-1.876	11.24
ΔSYNCH	32536	0.053	1.896	-0.625	0.047	0.725
ΔTURN	39863	-0.065	5.437	-1.535	-0.018	1.438

Panel B: Correlation matrix for level variables

This panel presents correlation coefficients for the levels of key variables in our sample. Variable definitions are provided in Appendix A. Pearson (Spearman) coefficients are presented above (below) the diagonal.

	ETF	HLSPREAD	ILLIQ_N	LN(MVE)	BTM	TURN	SYNCH
ETF	1.000	-0.141	-0.107	0.219	0.010	0.171	0.350
HLSPREAD	-0.129	1.000	0.896	-0.451	0.158	0.136	-0.113
ILLIQ_N	-0.110	0.908	1.000	-0.374	0.163	0.244	-0.051
LN(MVE)	0.326	-0.501	-0.417	1.000	-0.276	0.185	0.524
BTM	-0.010	0.102	0.080	-0.370	1.000	-0.076	-0.106
TURN	0.403	0.143	0.232	0.389	-0.230	1.000	0.158
SYNCH	0.448	-0.106	-0.047	0.533	-0.056	0.240	1.000

Panel C: Correlation matrix for change variables

This panel presents correlation coefficients for the changes of key variables in our sample. Variable definitions are provided in Appendix A. Pearson (Spearman) coefficients are presented above (below) the diagonal.

	ΔETF	ΔHLSPREAD	ΔILLIQ_N	ΔLN(MVE)	ΔBTM	ΔTURN	ΔSYNCH
ΔETF	1.000	0.171	0.193	-0.005	0.022	0.070	0.080
ΔHLSPREAD	0.166	1.000	0.841	-0.102	0.138	0.118	0.142
ΔILLIQ_N	0.177	0.790	1.000	-0.043	0.104	0.092	0.169
ΔLN(MVE)	-0.016	-0.033	0.021	1.000	-0.722	0.096	0.148
ΔBTM	0.026	0.043	-0.011	-0.792	1.000	-0.007	-0.114
ΔTURN	0.067	0.132	0.091	0.051	-0.004	1.000	0.083
ΔSYNCH	0.112	0.245	0.285	0.164	-0.121	0.078	1.000

Table 2: Regressions of bid-ask spread on ETF ownership

This table presents regression summary statistics from the below regression of changes in bid-ask spread (Δ*HLSPREAD*) on changes in ETF ownership (Δ*ETF*). *t*-statistics based on standard errors double clustered by firm and year are shown in parentheses. See Appendix A for variable descriptions. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, based on a two-sided test.

$$\Delta HLSPREAD_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \sum_k \beta_k \, \Delta Controls_{it-1} + \beta_{ind} + \beta_t + \epsilon_{it}$$

Y = ΔHLSPREAD$_t$

	Pred.	I		II	
ΔETF_{t-1}	+	0.016	**	0.017	**
		(2.41)		(2.51)	
$\Delta INST_{t-1}$				-0.0002	
				(-0.58)	
$\Delta LN(MVE)_{t-1}$		-0.044	*	-0.043	
		(-1.67)		(-1.64)	
ΔBTM_{t-1}		0.051	**	0.051	**
		(2.42)		(2.43)	
$\Delta TURN_{t-1}$		0.000		0.001	
		(0.55)		(0.61)	
$\Delta STD(RET)_{t-1}$		0.001	*	0.001	*
		(1.86)		(1.85)	
Year FE		YES		YES	
Industry FE		YES		YES	
N		39863		39863	
Adj. R-Square		0.304		0.304	

Table 3: Regressions of a measure of illiquidity (*ILLIQ_N*) on ETF ownership

This table presents regression summary statistics from the below regression of changes in daily average absolute returns ($\Delta ILLIQ\_N$) on changes in ETF ownership (ΔETF). *t*-statistics based on standard errors double clustered by firm and year are shown in parentheses. See Appendix A for variable descriptions. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, based on a two-sided test.

$$\Delta ILLIQ\_N_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \beta_3 \Delta ILLIQ\_D_{it} + \sum_k \beta_k \Delta Controls_{it-1} + \beta_{ind} + \beta_t + \epsilon_{it}$$

		Y = ΔILLIQ_N$_t$			
	Pred.	**I**		**II**	
ΔETF_{t-1}	+	0.042	**	0.043	**
		(2.55)		(2.57)	
$\Delta INST_{t-1}$				-0.0004	
				(-0.61)	
$\Delta LN(MVE)_{t-1}$		-0.015		-0.015	
		(-1.55)		(-1.55)	
ΔBTM_{t-1}		0.172	***	0.171	**
		(4.20)		(4.25)	
$\Delta ILLIQ\_D_t$		0.004	***	0.004	***
		(5.38)		(5.44)	
Year FE		YES		YES	
Industry FE		YES		YES	
N		39863		39863	
Adj. R-Square		0.371		0.371	

Table 4: Regressions of stock return synchronicity (*SYNCH*) on ETF ownership

This table presents regression summary statistics from the below regressions of changes in stock return synchronicity (Δ*SYNCH*) on changes in ETF ownership (Δ*ETF*). *t*-statistics based on standard errors double clustered by firm and year are shown in parentheses. See Appendix A for variable descriptions. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, based on a two-sided test.

$$\Delta SYNCH_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \sum_k \beta_k \, \Delta Controls_{it-1} + \beta_{ind} + \beta_t + \epsilon_{it}$$

Y = ΔSYNCH$_t$

	Pred.	I		II	
ΔETF_{t-1}	+	0.090	***	0.090	***
		(3.70)		(3.67)	
$\Delta INST_{t-1}$				0.0002	
				(0.21)	
$\Delta LN(MVE)_{t-1}$		0.595	***	0.593	***
		(8.78)		(9.82)	
ΔBTM_{t-1}		-0.093		-0.094	*
		(-1.62)		(-1.70)	
$\Delta TURN_{t-1}$		0.001		0.001	
		(1.11)		(1.12)	
$\Delta SKEW_{t-1}$		0.004		0.004	
		(0.41)		(0.41)	
$\Delta BETA_{t-1}$		-0.704	***	-0.704	***
		(-14.85)		(-14.58)	
Year FE		YES		YES	
Industry FE		YES		YES	
N		32536		32536	
Adj. R-Square		0.086		0.086	

Table 5: Regressions of current returns on future earnings and ETF ownership
Panel A: Total earnings

This table presents regression summary statistics of regressions of current annual stock returns (RET_t) on total future earnings ($EARN_{t+1}$) , the lagged level of ETF ownership (ETF_{t-1}) and lagged changes in ETF ownership (ΔETF_{t-1}). *t*-statistics based on standard errors double clustered by year and firm are shown in parentheses. See Appendix A for variable descriptions. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, based on a two-sided test.

Y = RET_t						
	Variable type	Pred	I		II	
$EARN_{t-1}$	Main effect		-0.435	***	-0.438	***
			(-6.90)		(-6.14)	
EARN	Main effect		0.282	***	0.284	***
			(3.76)		(3.81)	
$EARN_{t+1}$	Main effect		0.015	***	0.015	***
			(4.80)		(4.32)	
ETF_{t-1}	Main effect		-0.170			
			(-0.50)			
ΔETF_{t-1}	Main effect				-0.768	**
					(-2.06)	
ETF_{t-2}	Main effect				0.070	
					(0.17)	
$ETF_{t-1} \times EARN_{t-1}$	Interaction		0.594			
			(0.68)			
$ETF_{t-1} \times EARN_t$	**Interaction**	–	**-3.662**	***		
			(-2.60)			
$ETF_{t-1} \times EARN_{t+1}$	**Interaction**	–	**-0.212**	***		
			(-4.64)			
$\Delta ETF_{t-1} \times EARN_{t-1}$	Interaction				0.194	
					(0.37)	
$\Delta ETF_{t-1} \times EARN_t$	**Interaction**	–			**-3.636**	***
					(-2.68)	
$\Delta ETF_{t-1} \times EARN_{t+1}$	**Interaction**	–			**-0.195**	**
					(-2.07)	
$ETF_{t-2} \times EARN_{t-1}$	Interaction				0.797	
					(0.64)	
$ETF_{t-2} \times EARN_t$	**Interaction**	–			**-3.739**	**
					(-2.23)	
$ETF_{t-2} \times EARN_{t+1}$	**Interaction**	–			**-0.210**	***
					(-4.48)	

INST $_{t-1}$	Control		-0.157	***	-0.160	***
			(-3.74)		(-3.79)	
LN(MVE) $_{t-1}$	Control		0.011	*	0.010	*
			(1.87)		(1.82)	
LOSS $_{t}$	Control		-0.213	***	-0.212	***
			(-8.61)		(-8.61)	
ATGROWTH $_{t}$	Control		0.002	***	0.002	***
			(9.73)		(9.76)	
RET $_{t+1}$	Control		-0.026		-0.026	
			(-0.62)		(-0.61)	
Year FE			YES		YES	
Industry FE			YES		YES	
N			29970		29970	
Adj R-Square			0.350		0.351	

Panel B: Earnings components

This table presents regression summary statistics of regressions of current annual stock returns (RET_t) on future aggregate and firm-specific earnings ($EARNAGG_{t+1}$ and $EARNFIRM_{t+1}$), the lagged level of ETF ownership (ETF_{t-1}) and lagged changes in ETF ownership (ΔETF_{t-1}). Main effects variables include ΔETF_{t-1}; ETF_{t-1}; and lagged, current, and future levels of *EARNAGG* and *EARNFIRM*. Control variables include $INST_{t-1}$, $LN(MVE)_{t-1}$, $ATGROWTH_t$, $LOSS_t$, and RET_{t+1}. *t*-statistics based on standard errors double clustered by year and firm are shown in parentheses. See Appendix A for variable descriptions. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, based on a two-sided test.

Y = RET_t				
	Pred.	**I**	**II**	
$\Delta ETF_t \times EARNAGG_t$	+	0.868	2.765	***
		(0.69)	(4.25)	
$\Delta ETF_t \times EARNFIRM_t$		0.888	2.921	***
		(0.80)	(4.46)	
$\Delta ETF_{t-1} \times EARNAGG_{t-1}$			1.660	
			(1.40)	
$\Delta ETF_{t-1} \times EARNFIRM_{t-1}$			-0.577	
			(-0.66)	
$\Delta ETF_{t-1} \times EARNAGG_t$	–		-3.670	***
			(-2.72)	
$\Delta ETF_{t-1} \times EARNFIRM_t$	–		-1.874	
			(-1.57)	
$\Delta ETF_{t-1} \times EARNAGG_{t+1}$	–		-0.104	
			(-0.69)	
$\Delta ETF_{t-1} \times EARNFIRM_{t+1}$	–		-0.172	***
			(-3.26)	
$ETF_{t-2} \times EARNAGG_{t-1}$			2.708	
			(1.53)	
$ETF_{t-2} \times EARNFIRM_{t-1}$			0.286	
			(0.18)	
$ETF_{t-2} \times EARNAGG_t$	–		-6.387	**
			(-6.19)	
$ETF_{t-2} \times EARNFIRM_t$	–		-4.255	**
			(-4.16)	
$ETF_{t-2} \times EARNAGG_{t+1}$	–		-0.097	
			(-0.94)	
$ETF_{t-2} \times EARNFIRM_{t+1}$	–		-0.136	**
			(-2.02)	
Year and industry FE		YES	YES	
Main effects included		YES	YES	
Controls included		YES	YES	
N		29970	29970	
Adj. R-Square		0.373	0.384	

Table 6: Replication and reconciliation of Glosten et al. (2015; GNZ) results

This table presents summary statistics for regressions of quarter *t* returns on quarterly seasonally-adjusted earnings measures (*earn*) and quarterly ETF ownership changes (*etfchange*) from varying periods. *t*-statistics based on standard errors double clustered by year and firm are shown in parentheses. Control variables include an intercept, a loss indicator (SLOSS $_{t}$), lagged market-to-book ratio (MTB $_{t-1}$), lagged standard deviation of earnings (STD $_{t-1}$), lagged level of ETF ownership (ETF $_{t-1}$), lagged firm size (LN(MVE) $_{t-1}$), and the interaction of earnings with the loss indicator (*earn* × SLOSS). See Appendix A for variable descriptions. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, based on a two-sided test.

		Y = Quarter *t* returns						
		Varying ETF change period			**Varying earnings period**			
		I	**II**	**III**	**IV**	**V**	**VI**	**VII**
Alternative measures of *etfchange*	δETF_{t}	5.344*** (6.535)						
	δETF_{t-1}		0.083 (0.123)					
	$\delta ETF_{SUM\ t-1\ to\ t-4}$			-1.186* (-1.660)	-1.034 (-1.508)	-1.092 (-1.574)	-1.228* (-1.758)	-0.948 (-1.394)
Alternative measures of *earn*	$SEARN_{t}$	0.488** (2.322)	0.564** (2.436)	0.730*** (2.926)				
	$SEARN_{t+1}$				1.449*** (2.885)			
	$SEARN_{t+2}$					1.504*** (3.337)		
	$SEARN_{t+3}$						1.081*** (2.947)	
	$SEARN_{SUM\ t\ to\ t+3}$							0.526*** (3.089)
	earn* × *etfchange	**0.438**** **(2.458)**	**0.386*** **(1.729)**	**0.101** **(0.767)**	**-0.418** **(-1.494)**	**-0.376** **(-1.466)**	**-0.568*** **(-1.809)**	**-0.198*** **(-1.848)**
	Controls	YES	YES	YES	YES	YES	YES	YES
	Obs	106953	101495	101495	101483	101483	101483	101483
	Adj R-Square	0.0295	0.0282	0.0281	0.0407	0.0407	0.0193	0.0451

Table 7: Regressions of analyst coverage on ETF ownership

This table presents summary statistics for regressions of changes in analyst coverage ($\Delta ANALYST$) on changes in ETF ownership (ΔETF). *t*-statistics based on standard errors double clustered by firm and year are shown in parentheses. See Appendix A for variable descriptions. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, based on a two-sided test.

Y = Δ *ANALYST$_t$*	Pred.	**I**		**II**		**III**		**IV**	
ΔETF$_{t-1}$	–	**0.042**		**0.017**		**0.029**		**0.008**	
		(1.24)		**(0.46)**		**(0.87)**		**(0.23)**	
ETF$_{t-2}$	–					**-0.033**	***	**-0.025**	***
						(-4.52)		**(-3.25)**	
ΔINST$_{t-1}$				0.013	***			0.012	***
				(9.06)				(8.03)	
INST$_{t-2}$								0.000	
								(-0.35)	
LN(MVE)$_{t-1}$		0.056		0.060		0.065		0.066	
		(0.98)		(1.05)		(1.11)		(1.15)	
ΔBTM$_{t-1}$		-0.169	**	-0.143	*	-0.171	**	-0.146	*
		(-1.96)		(-1.72)		(-1.97)		(-1.74)	
ΔTURN$_{t-1}$		0.014	***	0.008	*	0.013	***	0.008	*
		(3.09)		(1.94)		(2.95)		(1.89)	
ΔSTD(RET)$_{t-1}$		0.000		0.001		0.000		0.001	
		(0.20)		(1.17)		(0.36)		(1.25)	
ΔINTAN_F$_{t-1}$		1.386	***	1.344	***	1.382	***	1.342	***
		(3.77)		(3.66)		(3.75)		(3.64)	
ΔRD_F$_{t-1}$		1.965	***	2.011	***	1.977	***	2.018	***
		(2.79)		(2.85)		(2.81)		(2.86)	
MOM$_{t-1}$		0.009	***	0.009	***	0.009	***	0.009	***
		(7.68)		(7.44)		(7.70)		(7.48)	
Year and industry FE		YES		YES		YES		YES	
N		29562		29562		29562		29562	
Adj. R-Square		0.038		0.041		0.039		0.042	