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

by

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

First Draft: September 12, 2014

Current Draft: July 26, 2015

Abstract

In a noisy rational expectations framework with costly information, some agents expend resources to become informed, and earn a return for their efforts by trading with the uninformed. Applying this insight, we examine the proposition that 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 (measured as bid-ask spreads and price impact of trades); (2) an increase in "stock return synchronicity" (measured as the co-movement of firm-level stock returns with general market and related-industry stock returns); (3) a decline in "future earnings response coefficients" (measured as the predictive power of current returns for future earnings), 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, a combination which results in less informative security prices for the component firms.

JEL Classifications: G11, G14, M41

Keywords: Exchange traded funds (ETFs); Uninformed and informed traders; Costly information; Trading costs; Pricing efficiency.

** 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 (suhas.sridharan@anderson.ucla.edu) is at the Anderson School of Management, UCLA. We gratefully acknowledge comments from seminar participants at UCLA, Emory University, Tel Aviv University, the Interdisciplinary Center (IDC) Herzliya, and research assistance from Woo Young Park and Padmasini Venkatachari.

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

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).

In this paper, we use a natural experiment 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 exchange-traded fund (ETF) ownership on the pricing 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 informativeness of its share price. However, as we argue below, market

---

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

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

frictions related to information acquisition costs can cause ownership by ETFs to be a significant economic event, with direct consequences for price informativeness in the underlying securities.

Our central conjecture is that ETF ownership can influence a stock's price informativeness 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 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. In addition, ETFs offer an attractive investment alternative for uninformed traders who would otherwise trade the underlying component securities.[3] These two 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

[3] Uninformed investors are drawn to the ETF market because their losses to informed traders are lower in this market than in the market for individual securities (e.g., Rubinstein 1989; Subrahmanyam 1991; Gorton and Pennacchi 1993).

information environment, and a reduction in the extent to which its stock price is able to quickly reflect firm-specific information.[4]

To test these hypotheses, we conduct a series of tests using a cross-section of U.S. stocks between 2000 and 2011. Our research design makes use of panel-data based on firm-quarter and firm-year observations. To conduct these tests, we first collect end-of-quarter ETF ownership data for all firms. We then examine the effect of both the *level* of ETF ownership and the *changes* 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 (e.g., 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*.[5] 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 either next quarter or next year.

For example, on average, High-ETF-Ownership firms, defined as firms in which ETF ownership equals or exceeds 3%, have average bid-ask spreads over the next quarter (or next year) that are approximately 6% higher than those of firms with low or no ETF ownership. These same High-ETF-Ownership firms exhibit average daily absolute equity returns that are approximately 3% higher than those of firms with low or no ETF ownership. The results are

[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 traders. In Section 2, we explain in detail the implications of this difference for our tests.

[5] 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*.

similar when we examine the association between quarterly or annual *changes* in ETF ownership and corresponding *changes* in both proxies of liquidity. In all these tests, both the level and the change in ETF ownership portend lower firm-level liquidity.[6]

Our second hypothesis deals with information-related effects of ETF ownership. Because we expect the information-related effects to take place in the longer run, we use the annual panel to test this hypothesis. Specifically, 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* (i.e., the extent to which firm-specific stock return variation is attributable to general market and related-industry movements), and (2) future earnings response coefficient, *FERC* (i.e., 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.[7]

Our results are broadly consistent 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*. For example, on average over the next year, High-ETF-Ownership firms exhibit stock return synchronicity levels that are 45% higher than those of firms with low or no ETF ownership. Similarly, on average, firms with high ETF ownership possess FERC coefficients that are 5% lower than those of firms with low or no ETF ownership. 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

[6] Hamm (2013) reports that increased ETF ownership is associated with an increase in the adverse selection component of the bid-ask spread. Our results are consistent with hers, although we use alternative measures of liquidity, different control variables, and a more complete firm-level longitudinal data set.
[7] 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).

prior literature (e.g., Roll 1988, Durnev et al. 2003, Piotroski and Roulstone 2004, Ettredge et al. 2005, Choi et al. 2011).

Finally, we find that an increase in ETF ownership is also accompanied by a decline in the number of analysts covering the firm. Specifically, after controlling for firm size, book-to-market ratio, return and earnings volatilities, share turnover, and levels of intangible assets and research and development expenses, we observe that firms with high ETF ownership are, on average, covered by one fewer analyst relative to firms with low or no ETF ownership.

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.[8] A number of prior studies 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). 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 (e.g., Ben-David et al. 2014, Broman 2013, Krause et al. 2013), particularly in times of market stress (Wurgler 2010).

Our study adds an informational perspective to this debate. Adopting key insights from information economics (e.g., Rubinstein 1989, Subrahmanyam 1991, Gorton and Pennacchi

[8] 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.

1993), 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 (e.g., 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 empirical support for a long-standing prediction of the noisy rational expectations literature. A number of models in this literature (e.g., Grossman and Stiglitz 1980, Admati 1985, Diamond and Verrecchia 1981, Verrecchia 1982) predict that when information is costly to acquire and process, the information efficiency of security prices will vary with the number of uninformed investors willing to trade the underlying 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 (2014) 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 highlight 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 3, we develop our main hypotheses and outline our research design. Section 4 reports the empirical findings, and Section 5 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.[9] In addition, ETFs do not sell shares directly to investors. Instead, they only issue shares in large blocks called "creation units" to authorized participants ("AP"s) 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 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.

[9] 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.

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, traders 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).

Essentially, ETFs offer the convenience of a stock (e.g., ETFs can be bought and sold during the day, like common stocks) along with the diversification of a mutual fund or index funds (i.e., 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 resemble many characteristics of what Rubinstein (1989) identifies as 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 the underlying basket and organization costs), (3) enhance tax benefits obtained from positions in the individual components of the basket (e.g., 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 the ETFs especially attractive to uninformed traders who would otherwise trade the underlying component securities.[10] 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 component 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 that represent the channel as well as the consequences of a change in a firm's stock pricing efficiency: (1) transactions costs of market participants; (2) the extent to which stock prices reflect firm-specific information. Accordingly, we 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 securities for uninformed traders. Because of the trading benefits offered by ETFs, especially to uninformed investors, we expect the uninformed investors to gravitate towards ETFs and away from the underlying stocks (e.g., Milgrom and Stokey 1982, Rubinstein 1989). As uninformed traders shift towards trading ETFs and away from trading the underlying securities, transactions costs

[10] 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.

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 propose and test 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 a deterioration in the pricing efficiency of the underlying securities.

To test H1, we examine the relation between ETF ownership and two proxies of liquidity that capture trading costs: (1) bid-ask spreads, and (2) price impact of trades (e.g., Goyenko et al. 2009). To investigate to the relation between ETF ownership and bid-ask spreads, we estimate several versions of the following regression model using quarterly and annual measures:[11]

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

The dependent variable in Eq. (1), $HLSPREAD_{it}$, is the Corwin and Schultz (2012) quarterly or annual high-low measure of bid-ask spread for firm *i* over quarter or year *t*. We use this measure of liquidity 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

[11] We test H1 using both quarterly and annual panels because we expect the liquidity effects of increased ETF ownership to exist in the short-run and in the long-run after an increase in ETF ownership.

(2009) techniques for measuring bid-ask spreads.[12] The variable of interest in Eq. (1), ETF_{it-1}, is the percentage of firm *i*'s shares held by all ETFs as of the end of quarter or year *t-1*. $INST_{it-1}$ is the percentage of firm *i*'s shares held by all institutions as of the end of quarter or year *t-1*,

ETF ownership is highly correlated with overall institutional ownership and prior research suggests there might be a relation between institutional ownership and bid-ask spreads.[13] To isolate the effect of 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 $INST_{it-1}$ directly in Eq. (1), as an additional control variable.[14]

In Eq. (1), $Controls_{it-1}$ represents a vector of firm and industry related control variables nominated by prior literature. We control for the log of market value of equity (*LN(MVE)*) as of the end of quarter or 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 (e.g., Copeland and Galai (1983)). Accordingly, we also control for the annualized standard deviation of daily returns during quarter or year *t-1* (*STD(RET)*) and average share turnover during quarter or year *t-1* (*TURN*). We also include the book to market ratio (*BTM*) at the end of quarter or 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-quarter or year and industry fixed effects.[15]

---

[12] Our inferences are the same when we employ alternative measures of bid-ask spread and use them as proxies for trading costs.

[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. Consistent with this notion, Agarwal (2011) documents a non-monotonic (U-shaped) relation between spreads and institutional ownership.

[14] Our inferences are the same when we use the residual from the regression model $ETF_{it} = \beta_0 + \beta_1 INST_{it} + \varepsilon_{it}$ as a measure of ETF ownership that is orthogonal to the level of institutional ownership.

[15] The industry fixed effects are defined based on the 48 Fama and French (1997) industry classification.

Our first hypothesis (H1) predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus,* firms with higher levels of ETF ownership experience higher bid-ask spreads.

Because the level of ETF ownership is not stationary over time, we also estimate versions of Eq. (1) using changes in the quarterly and annual ETF ownership measures:

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

In Eq. (1A), 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 quarter or year *t* and its value in quarter or year *t-1*. Similarly, ΔETF_{it-1} is the difference between firm *i*'s measure of *ETF* during quarter or year *t-1* and its value in quarter or year *t-2*. 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.

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 well-accepted proxy for price impact of trades (e.g., 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 effect 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 several versions of the following regression model using quarterly and annual measures:

$$ILLIQ\_N_{it} = \beta_1 ETF_{it-1} + \beta_2 INST_{it-1} + \beta_3 ILLIQ\_D_{it} + \sum_k \beta_k Controls_{it-1} + \sum_j \beta_j INDST\_FE_i + \sum_l \beta_l YEARQTR\_FE_t + \epsilon_{it} \quad (2)$$

The dependent variable in Eq. (2), $ILLIQ\_N_{it}$, is the daily absolute return for firm *i* averaged over all the trading days in quarter or year *t*. $ILLIQ\_D_{it}$ is the daily dollar volume for firm *i* averaged over all the trading days in quarter or year *t*. ETF_{it-1} and $INST_{it-1}$ are as defined above. $Controls_{it-1}$ denotes several control variables measured as of the end of quarter or year *t-1*. Specifically, it includes the log of market value of equity (*LN(MVE)*) as of the end of quarter or year *t-1*, because we expect larger firms to exhibit smaller price impact, and book-to market-ratio (*BTM*) at the end of quarter or year *t-1* to control for the effects of financial distress and/or growth opportunities. Our hypothesis predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus* (e.g., controlling for *ILLIQ_D*), higher levels of ETF ownership are associated with greater price impact of trades trading (and hence lower liquidity or higher trading costs for market participants).

Analogous to Eq. (1A), we also estimate versions of Eq. (2) using changes (as opposed to levels) in quarterly and annual measures:

$$\Delta ILLIQ\_N_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \beta_3 \Delta ILLIQ\_D_{it} + \sum_k \beta_k \Delta Controls_{it-1} + \sum_j \beta_j INDST\_FE_i$$

$$+ \sum_l \beta_l YEARQTR\_FE_t + \epsilon_{it} \quad (2A)$$

In Eq. (2A), the Δ operator indicates a change in the value of a particular variable. For example, $\Delta ILLIQ\_N_{it}$ ($\Delta ILLIQ\_N_{it}$) is the difference between firm *i*'s measure of $ILLIQ\_N$ ($ILLIQ\_D$) during quarter or year *t* and its value in quarter or year *t-1*. ΔETF_{it-1} and $\Delta INST_{it-1}$ are as defined above. Our first hypothesis predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus* (e.g., controlling for Δ*ILLIQ_D*) an increase in ETF ownership is associated with an increase in price impact of trades.

H2 states that an increase in ETF ownership is associated with deterioration in pricing efficiency of the underlying 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, following H2, we predict that stock return synchronicity and ETF ownership will be positively associated.

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*):[16]

$$RET_{id} = \alpha + \beta_1 MKTRET_d + \beta_2 MKTRET_{d-1} + \beta_3 INDRET_d + \beta_4 INDRET_{d-1} + \epsilon_{id} \quad (3)$$

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.*[17] 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 continuous measure of synchronicity (e.g., Piortoski and Roulstone 2004, Hutton et al. 2010, Crawford et al. 2012, Hutton et al. 2010)[18]: $SYNCH_{it} = log\left(\frac{R_{it}^2}{1-R_{it}^2}\right)$. 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 several

---

[16] We test H2 using annual panels because we expect the information-related effects of increased ETF ownership to manifest itself in stock prices gradually over time after an increase in ETF ownership.
[17] 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., Piotroksi 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.
[18] 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.

versions of the following regression model using annual panels:

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

In Eq. (4), $Controls_{it-1}$ indicates several annual measures that prior research suggests are associated with stock return synchronicity. In particular, following Jin and Myers (2006) we control for the skewness of firm *i*'s returns over year *t-1* (*SKEW*). In addition, since Li et al. (2014) show that synchronicity is often confounded with systematic risk, we include CAPM beta as a control for a firm's systematic risk. As additional controls, we include the log of market value of equity as of the end of year *t-1 (LN(MVE))*, book-to-market ratio as of the end of year *t-1 (BTM)*, average share turnover during year *t-1* (*TURN)*, and year and industry fixed effects. ETF_{it-1} and $INST_{it-1}$ are defined and measured as before. Our second hypothesis predicts that the coefficient β_1 is positive, indicating that, *ceteris paribus,* firms with higher levels of ETF ownership experience more synchronous stock returns.

Analogous to Eq. (1A) and (2A), we also estimate versions of Eq. (4) using changes (as opposed to levels) in annual measures:

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

In Eq. (4A), $\Delta SYNCH_{it}$ is the difference between firm *i*'s measure of $SYNCH$ during year *t* and its value in year *t-1*. $\Delta Controls_{it-1}$ denotes annual changes in the control variables, calculated as the differences between the values during year *t*-1 and *t*-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 the future earnings response coefficient, i.e., 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 using annual panels:[19]

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

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 lagged 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).

[19] Consistent with prior literature, we estimate Eq. (5) using levels specification only. This is mainly because, as opposed to other equations, the dependent variable in Eq. (5) is firm-level stock returns, which is stationary over time. It should be noted that our inferences are the same if we also estimate Eq. (5) using changes specifications.

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 future earnings is negative, indicating that *FERCs* are lower for firms with higher ETF ownership.

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_8 is negative, indicating that, *ceteris paribus,* firms with higher levels of ETF ownership possess lower levels of future earnings response coefficients.

We also test H2 by examining how ETF ownership relates to the number of analysts covering the firm during a year. Our hypothesis predicts that firms with high ETF ownership will experience a deterioration in pricing efficiency. We thus predict that firms with higher levels of ETF ownership experience lower levels of analyst coverage. To test this, we estimate several versions of the following regression model using an annual panel:

$$ANALYST_{it} = \beta_1 ETF_{it-1} + \beta_2 INST_{it-1} + \sum_k \beta_k \, Controls_{it-1}$$
$$+ \sum_j \beta_j \, INDST\_FE_i \; + \sum_l \beta_l \, YEARQTR\_FE_t + \epsilon_{it} \qquad (6)$$

In Eq. (6), $ANALYST_{it}$ is the number of unique analysts on I/B/E/S providing forecasts of firm *i*'s year *t* earnings. As before, $Controls_{it-1}$ represents a number of control variables, measured as of the end of 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 proportion of research and development expenses relative to total operating expenses (RD_F_{it-1}) and the proportion of intangible assets relative to total assets ($INTAN\_F_{it-1}$) as controls. Following Lang and Lundholm (1996), we also control for return volatility ($STD(RET)_{it-1}$). 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. (6) also includes controls for firm size (i.e., the natural logarithm of market value of equity) and book-to- market ratio. Our second hypothesis predicts that the coefficient β_1 is negative, indicating that, *ceteris paribus,* firms with higher levels of ETF ownership are covered by fewer security analysts.

Finally, we estimate versions of Eq. (6) using changes (as opposed to levels) in annual measures:

$$\Delta ANALYST_{it} = \beta_1 \Delta ETF_{it-1} + \beta_2 \Delta INST_{it-1} + \sum_k \beta_k \, Controls_{it-1}$$
$$+ \sum_j \beta_j \, INDST\_FE_i \; + \sum_l \beta_l \, YEARQTR\_FE_t + \epsilon_{it} \qquad (6A)$$

In Eq. (6A), $\Delta ANALYST_{it}$ is the difference between the number of analysts that cover firm *i* at the end of year *t* and the number in year *t-1*. $\Delta Controls_{it-1}$ denotes annual changes in the control variables, calculated as the differences between the values during year *t*-1 and *t*-2. Our second hypothesis predicts that the coefficient β_1 is negative, indicating that, *ceteris paribus,* increases in ETF ownership are associated with decreases in levels of analyst coverage.

One concern is that ETF ownership might not be large enough in most firms to produce the hypothesized effects. To evaluate this possibility, we isolate the potential effects of ETF ownership for firm-quarters or firm-years with particularly high levels of ETF ownership and large changes in ETF ownership. To do so, we define an indicator variable, *ETF_HI,* to equal 1 for firm-quarters or firm-years for which 3% or more of their shares are held by ETFs. Analogously, we define a second indicator variable, Δ*ETF_HI* , to equal 1 for firm-quarters or firm-years that experienced a 0.10% or a higher percentage increase in ETF holdings from period to period. We expect the coefficients on *ETF_HI* and on Δ*ETF_HI* to have the signs predicted by the three hypotheses.

To control for potential time-series correlation between firm-specific measures, we base our inferences from all equations on *t*-statistics calculated using standard errors clustered by firm. One exception to this is in Eq. (5), where the dependent variable is annual firm-level stock returns; for this equation we base our inferences on *t*-statistics calculated using standard errors clustered by year (Gow et al. 2010). Appendix A provides detailed descriptions of all the variables used in our empirical tests.

IV. Empirical Analyses

IV.1 Sample construction and descriptive statistics

We determine quarter-end and 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, for each ETF, we first obtain quarterly equity holdings from the Thomson Financial S12 database. Then, for every stock being held, we define *ETF* as the total number of shares held by any ETF divided by total number of shares outstanding in that quarter. We repeat this process for every firm-quarter between 2000 and 2014 to construct our quarterly and annual panels. 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-quarters or firm-years not held by any ETF 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 0.1% in 2000 to 7.0% in 2014.

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-quarter or firm-year observation must have information on stock price, 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 quarter. We restrict our analyses to firms with non-negative book-to-market ratios in every quarter of our sample period. This results in a sample of 184,327 firm-quarters 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 32,285 firm-year observations. The number of observations included in each regression varies according to data requirements.

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. The mean (median) percentage ETF ownership is 3.10% (2.24%). This is much lower than the level of institutional ownership, which has a mean (median) of 54.83% (58.23%). The distributional statistics of both ETF and institutional ownership in our sample are consistent with prior literature (e.g., Hamm 2013, Jiambalvo 2002). Nevertheless, as our hypotheses predict, 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 *ILLIQ_N* and *ILLIQ_D,* the two components of the Amihud (2002) illiqudity ratio, have notably different variances. *ILLIQ_N* is very narrowly distributed with a standard deviation of 1.42, while *ILLIQ_D* exhibits a significantly larger standard deviation of 139.78. 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 ETF ownership on *ILLIQ_N* controlling for *ILLIQ_D*.

Table 1, panel B presents Pearson and Spearman correlation coefficients between the key variables in our empirical analysis. In our sample, *ETF* exhibits a positive Pearson (Spearman) correlation of 0.562 (0.677) with *INST*. This is consistent with prior research documenting the strong relation between *ETF* and institutional ownership (e.g., Hamm 2013). *ETF* is also positively correlated with firm size (Pearson coef. = 0.411) and turnover (Pearson coef. = 0.106). Panel B reveals that *ETF* is negatively correlated with one of the proxies of trading costs, *HLSPREAD* (Pearson coef. = -0.177). However, it is important to note that this measure of correlation is confounded by the relation of other variables such as INST and *LN(MVE)* that are negatively correlated with *HLSPREAD* (for *INST*, Pearson coef. = -0.242; for *LN(MVE)* Pearson coef. = -0.358) and as stated above are positively correlated with *ETF*.

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

Tables 2 and 3 present regression summary statistics from the estimation of Eqs. (1), (1A), (2), and (2A), which are designed to test our first hypothesis using two measures of liquidity that capture trading costs and various model specifications. Panels A of both tables use quarterly samples and panels B employ the annual samples to conduct the analyses. In testing the first hypothesis we use both the quarterly and annual samples in order to create the most complete panel possible and because we expect the liquidity-related effects of ETF ownership to be reflected in stock prices already in the short-run (i.e., the next quarter) and persist in the long-run (i.e., the next year).

Column 1 of panel A reveals that the level of bid-ask spread, *HLSPREAD*, exhibits the expected relations with our control variables. *HLSPREAD* is decreasing in firm size (coef. = -0.126, *t*-stat. = -23.34), increasing in the BTM ratio (coef. = 0.042, *t*-stat. = 2.42), and increasing in return volatility (coef. = 0.016, *t*-stat = 2.78). According to column 1 of table 2, panel A the level of ETF ownership is negatively related to bid-ask spreads (coef. = -0.008, *t*-stat. = -4.43). However, it is important to note that this observed association is confounded by the relation between institutional ownership with *HLSPREAD* and with *ETF*. Controlling for *INST*, column 2 reveals that consistent with our first hypothesis, overall there exists a significantly positive association between *ETF* ownership and *HLSPREAD* (coef. = 0.004, *t*-statistic = 1.87). The coefficient on institutional ownership itself is negative and significant (coef. = -0.003, *t*-stat. = -

10.89). This pattern suggests that the negative coefficient on ETF ownership in column 1 is due to ETF ownership serving as a proxy for institutional ownership, when the latter is omitted from the regression.

More importantly, column 3 of panel A, table 2 reveals that the indicator variable, *ETF_HI*, which equals one for firm-quarters for which the level of ETF ownership equals or exceeds 3% percent, is significantly positive (coef. = 0.064, *t*-stat. = 6.31) while the coefficient on *ETF* becomes statistically indistinguishable from 0 (coef. = -0.002, *t*-stat. = -1.16). This evidence suggests that the positive association between *ETF* and *HLSPREAD* is driven by firms with particularly high levels ETF ownership. Economically, the estimated coefficient implies that firm-quarters with 3% or more of ETF ownership experience bid-ask spreads that are 6.4% higher than those with low or no ETF ownership. This finding supports our hypothesis that, ceteris paribus, the trading costs of market participants increase with ETF ownership.

Columns 4 – 6 of panel A, table 2 present summary statistics from the estimation of Eq. (1A), which is designed to provide a more robust test for the relation between ETF ownership and bid-ask spread, namely, using changes in *ETF* and changes in *HLSPREAD* (as opposed to levels of these measures as in Eq. (1)). Column 4 reveals that controlling for changes in institutional ownership, consistent with H1, Δ*ETF* is significantly positively related to Δ*HLSPREAD* (coef. = 0.003, t-stat. = 1.93). Column 6 shows that the positive relation between Δ*ETF* and Δ*HLSPREAD* in column 5 is driven by high changes in ETF ownership, Δ*ETF_HI*. Δ*ETF_HI* is defined as an indicator variable which equals one for firm-quarters that experienced an increase of 0.1% or more in ETF ownership during the last quarter. The coefficient on this indicator variable is positive and significant (coef. = 0.008, *t*-statistic = 3.44), indicating that, on average, a 0.1% change in *ETF* is associated with 0.8% change in *HLSPREAD*.

Panel B of table 2, which presents summary statistics from the estimation of Eqs. (1) and (1A) using annual sample, reveals evidence that is consistent with that presented in panel A of table 2 using quarterly sample. In particular, columns 3 and 6 show that both the coefficient estimate on *ETF_HI* and on Δ*ETF_HI* are significantly positive (coef. = 0.059, *t*-stat. = 5.87; coef. = 0.016, *t*-stat. = 3.09), suggesting that firm-years with high ETF ownership (high changes in ETF ownership) over the previous year experience bid-ask spreads which are 5.9% (1.6%) higher than firm-years with no or low ETF ownership (no or low change in ETF ownership).

Results from estimating Eq. (2) and (2A), which use another proxy for liquidity that capture trading costs, provide additional evidence on the association between measures of ETF ownership and *ILLIQ_N*, which is consistent with our first hypothesis. For example, column 3 of panel A, table 3 suggests that controlling for *INST* (and *ILLIQ_D*) firm-quarters with high ETF ownership experience over the course of the next quarter measures of *ILLIQ_N* that are 3% higher than those with low or no ETF ownership (coef. = 0.03, *t*-stat. = 2.06). Similarly, column 3 of panel B, table 3 reveals that in an annual sample, firm-years with high ETF ownership experience over the course of the next year measures of *ILLIQ_N* that are 2.9% higher than those with low or no ETF ownership. Investigation of the relation between Δ*ETF* and Δ*ILLIQ_N* reveal a consistently positive association within both quarterly and annual samples (panel A of table 3 – coef. range from 0.006 to 0.008, *t*-stat. range from 1.68 to 3.15; panel B of table 3 – coef. range from 0.04 to 0.042, *t*-stat. range from 10.26 to 13.26).

Taken together, the results presented in tables 2 and 3 provide strong evidence in support for our first hypothesis that an increase in ETF ownership is accompanied by an increase in trading costs for market participants.

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

Tables 4 and 5 present regression summary statistics from the estimation of Eqs. (4), (4A), 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 Eqs. (4) and (4A), which are designed to examine the relation between ETF ownership and annual measures of stock return synchronicity, using both levels and changes specifications. Column 1 reveals that the levels of synchronicity, *SYNCH*, exhibit the expected relations with our control variables. Consistent with prior research (Li et al. 2014), *SYNCH* is increasing in firm size (coef. = 0.495, *t*-stat. = 62.66), and our measure of systematic risk, *BETA* (coef. = 2.051, *t*-stat. = 72.41). Columns 1-3 of table 4 reveal that the level of ETF ownership is significantly positively related to stock return synchronicity (coef. range from 0.058 to 0.095, *t*-stat. range from 7.88 to 12.07). Column 3 shows that the coefficient estimate on *ETF_HI* is significantly positive (coef. = 0.202, *t*-stat. = 7.13), implying that firm-years with high levels of ETF ownership experience stock-return synchronicity over the course of the next year that is 45% higher than firms with low or no ETF ownership.

In columns 4 through 6 of table 4 we present summary statistics from the estimation of Eq. (4A) which is designed to examine the relation between changes in ETF ownership and changes in stock return synchronicity. Columns 4 through 6 provide evidence consistent with changes in ETF ownership are accompanied by positive changes in stock return synchronicity over the course of the next year (coefficient estimates on Δ*ETF* range from 0.048 to 0.085, *t*-stat. range from 5.32 to 12.69). Notably, under the changes specification, the relation between Δ*ETF* and Δ*SYNCH* is markedly different from the relation between Δ*INST* and Δ*SYNCH*. As opposed to columns 2 and 3, columns 5 and 6 show that the relation between Δ*INST* and Δ*SYNCH* is negative, although not significant (*t*-stat. = -1.56 and -1.51). Importantly, column 6 of table 4 shows that the positive relation between Δ*ETF* and Δ*SYNCH* is driven primarily by firm-years with particularly large changes in ETF ownership, Δ*ETF_HI* (coef. = 0.149, *t*-stat. = 5.45).

Table 5 presents regression summary statistics from the estimation of Eq. (5), which is designed to examine the relation between ETF ownership and the extent to which current firm-level returns reflect future firm-specific earnings. Consistent with prior literature, we observe a positive future earnings response coefficient across columns 1 to 3 (0.054 to 0.045, *t*-statistics = 2.14 to 1.45). Column 2 reveals that the interaction of future earnings with ETF ownership carries a significantly negative coefficient (coef. = -0.663, *t*-stat. = -2.13). This suggests that controlling for *INST* and a host of other variables prescribed by prior literature (e.g., *LOSS*, *ATGROWTH*) 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 firm-specific earnings-related information.[20] Importantly, column 3 of table 5 reveals that the coefficient estimate on the interaction between *ETF_HI* and *RET*(+1) is significantly negative (coef. = -0.05, *t*-stat. -3.74), indicating that stock returns of firms with high levels of *ETF* ownership are substantially less reflective of future firm-specific earnings information than stock returns of firms with low or no ETF ownership.

Taken together, the results presented in tables 4 and 5 indicate that an increase in ETF ownership is associated with an 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

[20] Columns 2 and 3 of table 5 also reveal that while consistent with Jiambalvo et al. (2002) the coefficient estimates on the interaction of *INST* with future earnings are positive (coef = 0.002 and 0.025), they are not statistically significant (*t*-stat. = 0.05 and 0.62).

hypothesis that stock prices of firms with high ETF ownership are impounding less firm-specific information.

Table 6 presents regression summary statistics from the estimation of Eqs. (6) and (6A), which are also designed to test our second hypothesis. Column 1 reveals that the level of analyst coverage, *ANALYST*, exhibits the expected relations with our control variables (e.g., Barth et al. 2001). *ANALYST* is higher among firms with high levels of research and development activities, *RD_F* (coef. = 2.242, *t*-stat. = 2.96) and is increasing in stock return volatility, *STD(RET)* (coef. = 0.002, *t*-stat. = 2.26), consistent with the notion that demand for firm-specific information among such firms is higher.

In column I the coefficient on *ETF* is -0.161 (*t*-stat. = -5.26). The magnitude of this coefficient suggests that for each percentage point increase in ETF ownership, analyst coverage (defined as number of analysts) declines by .16. Alternatively, every 6.25 percent increase in ETF ownership would be associated with a firm losing one analyst. This is the opposite of the relation we observe between institutional ownership and analyst coverage. *INST* exhibits a positive coefficient (0.014) that is also significantly different from zero (*t*-stat. = 4.71), suggesting that firms with higher levels of institutional ownership have more analysts covering them. We again observe that the relation between ETF ownership and analyst coverage is concentrated amongst firms with high levels of ETF ownership; column 2 of table 5 reveals that the coefficient on *ETF_HI* is negative (-0.541) and highly significant (*t*-statistic = -4.12).

In columns 3 and 4 of table 6 we examine the relation between changes in ETF ownership and changes in analyst coverage. Column 3 reveals that on average, changes in ETF ownership and changes in analyst coverage do not have a significant relation. Δ*ETF* has a coefficient of 0.020 that is not significantly different from zero. However, we do observe the predicted negative relation when firms experience a particularly large change in ETF ownership. In column 4 we observe a significantly negative coefficient (-0.094) on Δ*ETF_HI.* The coefficient estimate on Δ*ETF* in column 4 is significantly positive (coef. = 0.042, *t*-stat. = 2.57), but the magnitude of this coefficient is not large enough to dominate the negative relation captured by the coefficient on Δ*ETF_HI*.[21]

One possible concern arising from our analysis of changes in analyst coverage (i.e., columns 3 and 4 of table 6) is that there is not sufficient variation in the dependent variable, Δ*ANALYST*, to justify the estimation of Eq. (6A) using the method of ordinary least squares. To ensure that our inferences are not driven by misspecification of the method needed to estimate the equation, we re-estimate Eq. (6A) treating it as a probit model. The results of this estimation are presented in columns 5 and 6 of table 6. Our findings are wholly consistent with those drawn using the results of columns 3 and 4. Overall, using analyst coverage as a proxy for the availability of firm-specific information, the results in table 6 suggest that as an increase in ETF ownership is associated with a reduction in the pricing efficiency of the underlying securities.

V. Summary and concluding remarks

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. Hence, the availability of uninformed traders is essential to the pricing efficiency of securities.

[21] Using the coefficient estimates in column 4 of table 5, we estimate that any change in ETF ownership more than 10 basis points but less than 1.8 percent would result, *ceteris paribus,* in a reduction in analyst coverage. From our univariate analyses presented in Panel A of Table 1, we know that such a large change almost never occurs.

In this paper, we use a natural experiment 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 exchange-traded fund (ETF) ownership on the pricing efficiency of the individual component securities underlying the fund.

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. Specifically, we hypothesize and find that increased ETF ownership is accompanied by a deterioration of the pricing efficiency of the underlying component securities.

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 higher levels of ETF ownership are associated with an increase in stock return synchronicity and a reduction in future earnings response coefficients. We also observe a negative association between ETF ownership and the number of analysts covering the firm. Collectively, the evidence presented in this paper suggests increased ETF ownership can lead to weaker information environments for the underlying firms.

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.

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.

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.

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., 2013. The Effects of ETFs on Stock Liquidity. Working paper. Ohio State University, December 23.

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.

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.

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. 2014. *Alphanomics: The Informational Underpinnings of Market Efficiency*, in the *Foundations and Trends in Accounting* Series. Now Publishers: forthcoming

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.

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.

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

Panel A: Variables used in quarterly panels

Variable:		Definition
ETF_{it}	=	The percentage of firm *i's* common shares outstanding held by ETFs at the end of quarter *t*
$INST_{it}$	=	The percentage of firm *i's* common shares outstanding held by institutions at the end of quarter *t*
ETF_HI_{it}	=	Indicator variable equaling one if more than three percent of firm *i*'s common shares outstanding are held by ETFs
$HLSPREAD_{it}$	=	The Corwin and Schultz (2012) measure of bid-ask spread for firm *i* in quarter *t*
$ILLIQ\_N_{it}$	=	The average over quarter *t* of absolute daily equity returns for firm *i* (the numerator of the Amihud (2002) illiquidity ratio)
$ILLIQ\_D_{it}$	=	The average over quarter *t* of dollar volume for firm *i* (the denominator of the Amihud (2002) illiquidity ratio)
$LN(MVE)_{it}$	=	The natural logarithm of firm *i*'s market value of equity at the end of quarter *t*
BTM_{it}	=	The book to market ratio of firm *i* at the end of quarter *t*
$TURN_{it}$	=	The ratio of the average number of firm *i*'s shares traded in quarter *t* to firm *i*'s total common shares outstanding in quarter *t*
$STD(RET)_{it}$	=	The standard deviation of firm *i*'s daily returns over quarter *t*
Δ	=	The change operator

Panel B: Variables used in annual panels

Variable:		Definition
ETF_{it}	=	The percentage of firm *i's* common shares outstanding held by ETFs at the end of year *t*
$INST_{it}$	=	The percentage of firm *i's* common shares outstanding held by institutions at the end of year *t*
ETF_HI_{it}	=	Indicator variable equaling one if more than three percent of firm *i*'s common shares outstanding are held by ETFs
$HLSPREAD_{it}$	=	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)
$ANALYST_{it}$	=	The number of unique analysts in I/B/E/S that provide forecasts of year *t* earnings for firm *i*
$LN(MVE)_{it}$	=	The natural logarithm of firm *i*'s market value of equity at the end of year *t*
BTM_{it}	=	The book to market ratio of firm *i* at the end of year *t*
$TURN_{it}$	=	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*
$STD(RET)_{it}$	=	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.
$BETA_{it}$	=	The β_1 coefficient from the firm-quarter estimation of the model $$(RET - R_f)_{id} = \alpha + \beta_1(MKT - R_f)_d + \epsilon_{id}$$ 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.
$SKEW_{it}$	=	The skewness of firm *i*'s daily returns over year *t*

$EARN_{it}$	=	Earnings before extraordinary items for firm *i* in year *t*
$ATGROWTH_{it}$	=	Growth in assets for firm *i* from year *t-1* to year *t*.
$LOSS_{it}$	=	Indicator variable equaling one if firm *i* experienced a loss (defined as *EARN* < *0*) in year *t*
RET_{it}	=	The annual return for firm *i* in year *t*
$INTAN\_F_{it}$	=	The ratio of intangible assets to total assets for firm *i* in year *t*
RD_F_{it}	=	The ratio of research and development expenses to total operating expenses assets for firm *i* in year *t*
Δ	=	The change operator

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 Appendix A.

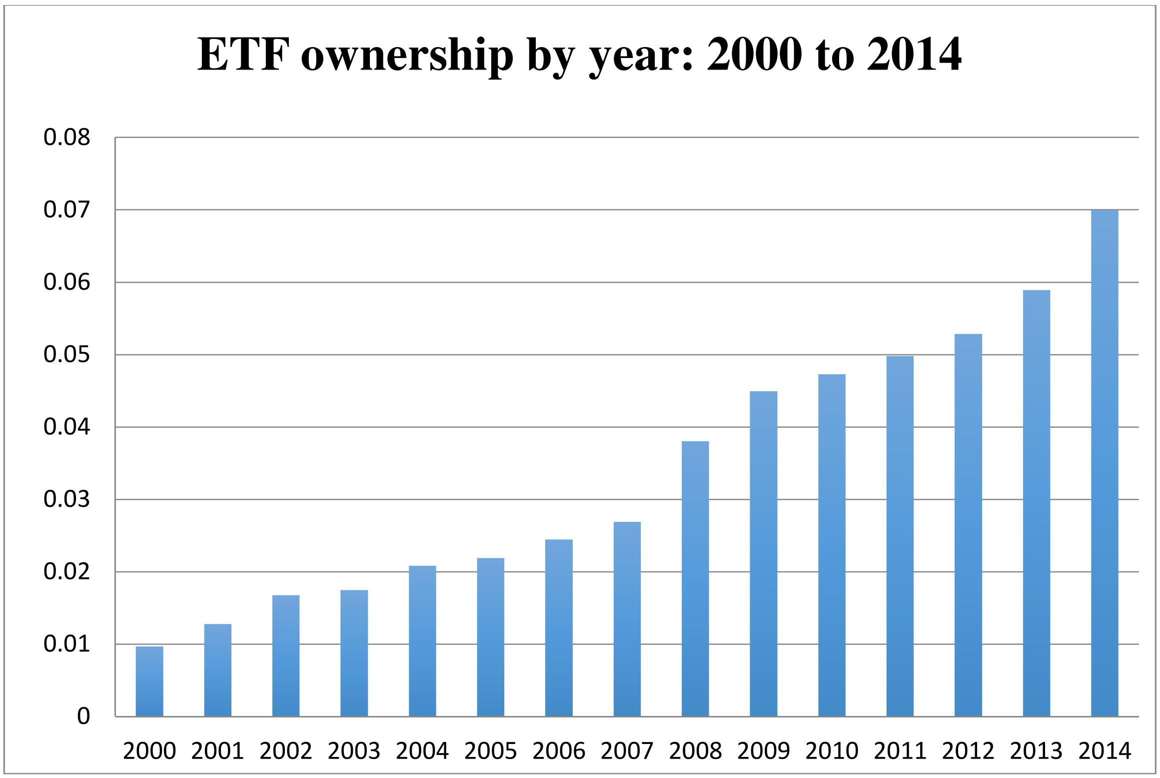
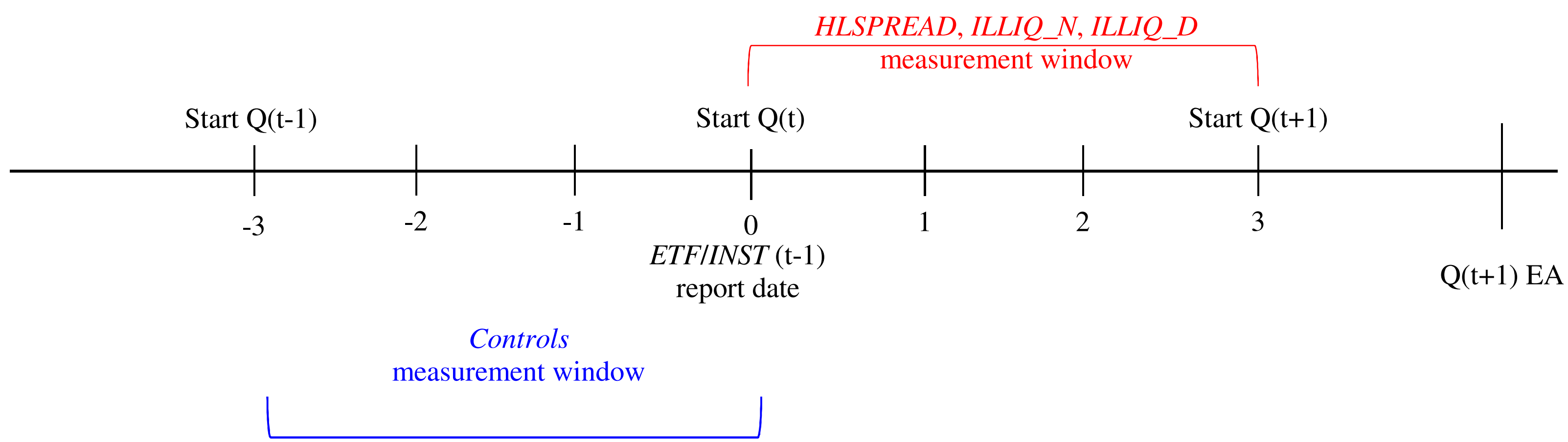


Figure 2: Sample construction timeline

Panel A: Quarterly timeline



Panel B: Annual timeline

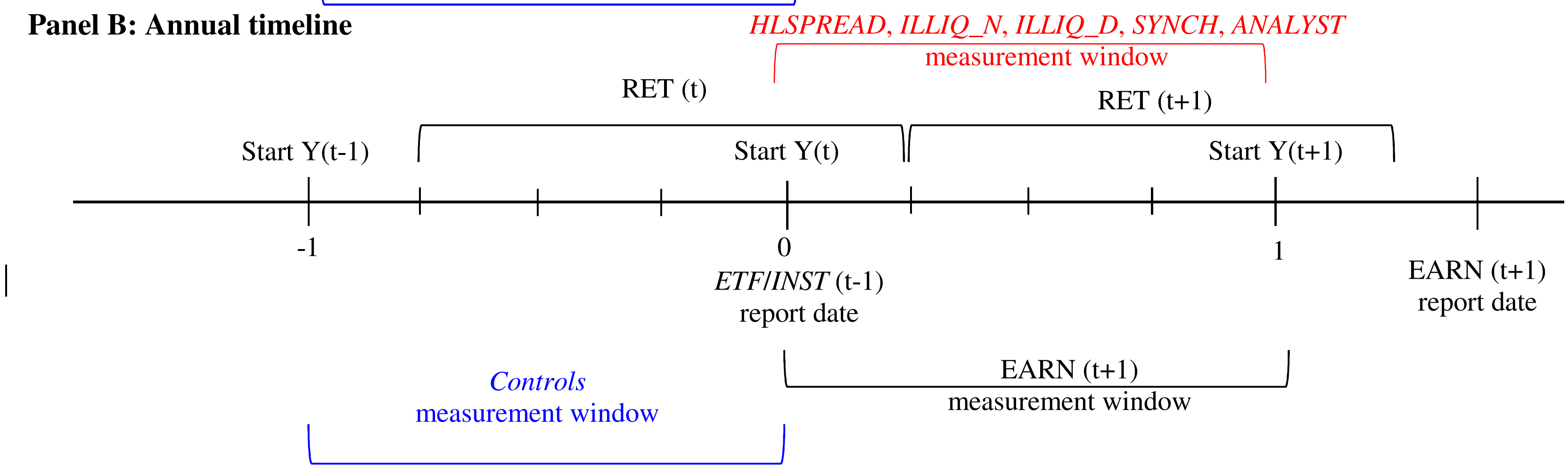


Table 1: Sample Description

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	StdDev	Q1	Median	Q3
INST	184327	54.83	32.01	26.13	58.23	83.53
ETF	184327	3.10	2.99	0.98	2.24	4.45
HLSPREAD	184327	1.13	0.87	0.60	0.90	1.39
LN(MVE)	184327	12.90	2.02	11.41	12.80	14.23
LN(TA)	184327	6.37	2.04	4.95	6.29	7.64
BTM	184327	0.75	0.94	0.34	0.57	0.90
TURN	184327	7.25	17.83	1.80	4.47	8.83
ILLIQ_N	182278	2.23	1.42	1.28	1.85	2.72
ILLIQ_D	182278	26.86	139.78	0.17	1.77	12.03
ETF_HI	184327	0.39	0.49	0.00	0.00	1.00
ΔETF	181913	0.12	0.61	-0.01	0.02	0.23
ANALYST	32285	7.73	6.88	2	6	11
SYNCH	37914	-2.29	2.51	-3.02	-1.53	-0.68

Panel B: Correlation matrix

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

	INST	ETF	HLSPREAD	LN(MVE)	BTM	TURN	STD(RET)
INST	1.000	0.677	-0.198	0.652	-0.196	0.599	-0.207
ETF	0.562	1.000	-0.180	0.565	-0.127	0.495	-0.247
HLSPREAD	-0.242	-0.177	1.000	-0.372	0.065	0.111	0.688
LN(MVE)	0.623	0.411	-0.358	1.000	-0.429	0.539	-0.367
BTM	-0.107	-0.065	0.165	-0.303	1.000	-0.326	0.082
TURN	0.163	0.106	0.059	0.124	-0.054	1.000	0.138
STD(RET)	-0.129	-0.108	0.310	-0.182	0.080	0.348	1.000

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

This table presents regression summary statistics from regressions of bid-ask spread (*HLSPREAD*) on ETF ownership (*ETF*). *t*-statistics based on standard errors clustered by 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.

Panel A: Quarterly panel

	LEVELS				CHANGES		
	I	II	III		IV	V	VI
ETF	-0.008***	0.004*	-0.002	ΔETF	0.003	0.003*	-0.000
	(-4.43)	(1.87)	(-1.16)		(1.53)	(1.93)	(-0.09)
INST		-0.003***	-0.003***	ΔINST		-0.000**	-0.000**
		(-10.89)	(-11.36)			(-2.43)	(-2.46)
ETF_HI			0.064***	ΔETF_HI			0.008***
			(6.31)				(3.44)
LN(MVE)	-0.126***	-0.105***	-0.106***	ΔLN(MVE)	-0.125***	-0.124***	-0.124***
	(-23.34)	(-21.42)	(-21.46)		(-14.50)	(-14.29)	(-14.31)
BTM	0.042**	0.049***	0.049***	ΔBTM	0.062***	0.061***	0.061***
	(2.42)	(2.62)	(2.61)		(5.64)	(5.56)	(5.53)
TURN	0.001	0.001	0.001	ΔTURN	0.002***	0.001***	0.001***
	(0.51)	(0.73)	(0.72)		(5.63)	(5.46)	(5.46)
STD(RET)	0.016***	0.016***	0.016***	ΔSTD(RET)	0.007***	0.007***	0.007***
	(2.78)	(2.76)	(2.76)		(12.04)	(11.97)	(11.97)
Industry FE	Yes	Yes	Yes	Industry FE	Yes	Yes	Yes
Year-Qtr FE	Yes	Yes	Yes	Year-Qtr FE	Yes	Yes	Yes
Observations	184327	184327	184327	Obs	181913	180895	180895
R-square	0.345	0.349	0.349	R-square	0.170	0.171	0.171

Panel B: Annual panel

	LEVELS				CHANGES		
	I	II	III		IV	V	VI
ETF	-0.004***	0.006***	0.000	ΔETF	0.015***	0.017***	0.012***
	(-2.87)	(3.16)	(0.27)		(8.95)	(9.69)	(5.65)
INST		-0.003***	-0.003***	ΔINST		-0.000	-0.000
		(-10.54)	(-10.98)			(-1.50)	(-1.52)
ETF_HI			0.059***	ΔETF_HI			0.016***
			(5.87)				(3.09)
LN(MVE)	-0.105***	-0.087***	-0.087***	ΔLN(MVE)	-0.060***	-0.061***	-0.061***
	(-20.85)	(-19.99)	(-20.07)		(-7.46)	(-7.43)	(-7.45)
BTM	0.020*	0.027***	0.028***	ΔBTM	0.039***	0.043***	0.043***
	(1.92)	(2.61)	(2.66)		(3.71)	(4.09)	(4.06)
TURN	-0.001	-0.001	-0.001	ΔTURN	0.000	0.000	0.000
	(-0.80)	(-0.40)	(-0.42)		(0.18)	(0.40)	(0.40)
STD(RET)	0.046***	0.045***	0.045***	ΔSTD(RET)	0.001***	0.001***	0.001***
	(7.85)	(7.77)	(7.77)		(8.90)	(9.42)	(9.43)
Industry FE	Yes	Yes	Yes	Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Year FE	Yes	Yes	Yes
Obs	46004	46004	46004	Obs	42634	41719	41719
R-square	0.391	0.395	0.396	R-square	0.282	0.288	0.289

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

This table presents regression summary statistics from regressions of the daily average absolute returns (*ILLIQ_N*) on ETF ownership (*ETF*). *t*-statistics based on standard errors clustered by 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.

Panel A: Quarterly panel

	LEVELS		
	I	II	III
ETF	-0.008***	0.001	-0.002
	(-3.18)	(0.19)	(-0.80)
INST		-0.002***	-0.002***
		(-6.58)	(-6.63)
ETF_HI			0.030**
			(2.06)
LN(MVE)	-0.238***	-0.214***	-0.214***
	(-37.09)	(-29.40)	(-29.36)
BTM	0.071***	0.079***	0.080***
	(3.90)	(4.06)	(4.05)
ILLIQ_D	0.002***	0.002***	0.002***
	(12.87)	(11.24)	(11.19)
Industry FE	Yes	Yes	Yes
Year-Qtr FE	Yes	Yes	Yes
Obs	182138	182138	182138
R-square	0.453	0.454	0.454

	CHANGES		
	IV	V	VI
ΔETF	0.011***	0.006*	0.008*
	(3.15)	(1.68)	(1.90)
ΔINST		0.006***	0.006***
		(14.80)	(14.80)
ΔETF_HI			-0.004
			(-0.97)
ΔLN(MVE)	-0.004***	-0.004***	-0.004***
	(-5.87)	(-6.08)	(-5.62)
ΔBTM	0.418***	0.427***	0.428***
	(25.55)	(26.06)	(26.07)
ΔILLIQ_D	0.000*	0.000	0.000
	(1.88)	(1.53)	(1.53)
Industry FE	Yes	Yes	Yes
Year Qtr FE	Yes	Yes	Yes
Obs	178247	177203	177203
R-square	0.278	0.281	0.281

Panel B: Annual panel

	LEVELS				CHANGES		
	I	II	III		IV	V	VI
ETF	-0.003	0.009***	0.006**	ΔETF	0.040***	0.042***	0.040***
	(-1.14)	(3.26)	(2.30)		(13.21)	(13.38)	(10.26)
INST		-0.003***	-0.003***	ΔINST		-0.000	-0.000
		(-8.95)	(-8.97)			(-1.24)	(-1.24)
ETF_HI			0.029*	ΔETF_HI			0.009
			(1.89)				(0.99)
LN(MVE)	-0.219***	-0.188***	-0.188***	ΔLN(MVE)	-0.015***	-0.015***	-0.015***
	(-34.48)	(-26.13)	(-26.10)		(-10.27)	(-10.50)	(-10.35)
BTM	0.055***	0.067***	0.067***	ΔBTM	0.169***	0.172***	0.171***
	(4.45)	(4.85)	(4.85)		(13.57)	(13.51)	(13.50)
ILLIQ_D	0.002***	0.002***	0.002***	ΔILLIQ_D	0.004***	0.004***	0.004***
	(13.62)	(11.53)	(11.49)		(15.92)	(15.70)	(15.67)
Industry FE	Yes	Yes	Yes	Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Year FE	Yes	Yes	Yes
Obs	43896	43896	43896	Obs	40808	39911	39911
R-square	0.452	0.455	0.455	R-square	0.366	0.371	0.371

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

This table presents regression summary statistics from regressions of stock return synchronicity (*SYNCH*) on ETF ownership (*ETF*). *t*-statistics based on standard errors clustered by 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.

	LEVELS				CHANGES		
	I	II	III		IV	V	VI
ETF	0.095***	0.074***	0.058***	ΔETF	0.085***	0.087***	0.048***
	(12.07)	(9.81)	(7.88)		(12.69)	(12.66)	(5.32)
INST		0.005***	0.005***	ΔINST		-0.001	-0.001
		(8.94)	(8.35)			(-1.56)	(-1.51)
ETF_HI			0.202***	ΔETF_HI			0.149***
			(7.13)				(5.45)
LN_MVE	0.495***	0.459***	0.458***	ΔLN(MVE)	0.541***	0.545***	0.546***
	(62.66)	(57.12)	(57.34)		(15.28)	(15.08)	(15.11)
BTM	-0.015	-0.029*	-0.027	ΔBTM	-0.116**	-0.108**	-0.109**
	(-0.91)	(-1.70)	(-1.59)		(-2.27)	(-2.09)	(-2.11)
TURN	-0.002	-0.002	-0.002	ΔTURN	0.022***	0.022***	0.022***
	(-1.32)	(-1.32)	(-1.32)		(11.20)	(11.35)	(11.38)
SKEW	-0.105***	-0.102***	-0.102***	ΔSKEW	0.007	0.006	0.006
	(-14.89)	(-14.35)	(-14.33)		(0.73)	(0.67)	(0.66)
BETA	2.051***	2.021***	2.008***	ΔBETA	-0.730***	-0.722***	-0.729***
	(72.41)	(72.79)	(72.78)		(-26.41)	(-25.82)	(-25.98)
Industry FE	Yes	Yes	Yes	Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Year FE	Yes	Yes	Yes
Obs	37900	37900	37900	Obs	33073	32523	32523
R-square	0.667	0.669	0.669	R-square	0.091	0.091	0.091

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

This table presents regression summary statistics of regressions of current annual stock returns (*RET*) on future earnings (*EARN(+1)*) and lag of ETF ownership (*ETF*). *t*-statistics based on standard errors clustered by year are shown in parentheses. As "additional controls" we include *ATGROWTH, LOSS, RET(+1),* and *LN(MVE).* 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.

	Pred.	I	II	III
		Estimate	Estimate	Estimate
EARN(-1)		0.000	-0.003	-0.001
		(-1.51)	(-1.64)	(-1.61)
EARN	+	0.009	0.060	0.058
		(0.44)	(1.16)	(1.17)
EARN(+1)	+	0.054**	0.051	0.045
		(2.14)	(1.55)	(1.45)
ETF		-1.321***	-0.093	-0.116
		(-4.45)	(-0.23)	(-0.37)
ETF_HI				0.001
				(0.03)
INST			-0.258***	-0.256***
			(-4.50)	(-4.10)
LOSS		-0.248***	-0.245***	-0.244***
		(-7.23)	(-7.27)	(-7.16)
ATGROWTH		0.000	0.000	0.000
		(1.56)	(1.53)	(1.53)
RET(+1)		-0.026	-0.023	-0.023
		(-0.70)	(-0.64)	(-0.64)
LN(MVE)		0.007	0.026***	0.026***
		(0.76)	(4.05)	(3.99)
ETFxEARN(-1)		0.012	-0.024	
		(1.59)	(-0.78)	
ETFxEARN		-0.872	0.008	
		(-0.34)	(0.00)	
ETFxEARN(+1)	–	-0.681	-0.663**	
		(-1.11)	(-2.13)	
ETF_HIxEARN(-1)				0.000
				(0.59)
ETF_HIxEARN				-0.028
				(-0.34)
ETF_HIxEARN(+1)	–			-0.050***
				(-3.74)

INSTxEARN(-1)		0.004	0.001
		(1.46)	(1.62)
INSTxEARN		-0.174	-0.167
		(-1.40)	(-1.13)
INSTxEARN(+1)		0.002	0.025
		(0.05)	(0.62)
INDST_FE	YES	YES	YES
YEAR_FE	YES	YES	YES
Adj.R	0.25	0.25	0.26
N	37517	37517	37517

Table 6: Regressions of analyst coverage on ETF ownership

This table presents regression summary statistics of regressions of analyst coverage (*ANALYST*) on ETF ownership (*ETF*). *t*-statistics based on standard errors clustered by 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.

Panel A: Annual panel

LEVELS			CHANGES	OLS		CHANGES	PROBIT	
	I	II		III	IV		V	VI
ETF	-0.161***	-0.117***	ΔETF	0.020	0.042**	ΔETF	0.008	0.018**
	(-5.26)	(-3.59)		(1.47)	(2.57)		(1.49)	(2.63)
INST	0.014***	0.016***	ΔINST	0.009***	0.009***	ΔINST	.004***	0.004***
	(4.71)	(5.47)		(6.46)	(6.43)		(6.59)	(6.56)
ETF_HI		-0.541***	ΔETF_HI		-0.094**	ΔETF_HI		-0.041**
		(-4.12)			(-2.47)			(-2.54)
LN(MVE)	3.009***	3.013***	ΔLN(MVE)	0.848***	0.848***	ΔLN(MVE)	0.369***	0.368***
	(54.50)	(54.63)		(14.42)	(14.41)		(15.32)	(15.31)
BTM	0.322***	0.312***	ΔBTM	0.421***	0.420***	ΔBTM	0.176**	0.176***
	(4.30)	(4.37)		(6.89)	(6.88)		(6.92)	(6.91)
STD(RET)	0.002**	0.002**	ΔSTD(RET)	0.002***	0.002***	ΔSTD(RET)	0.001***	0.001***
	(2.26)	(2.24)		(3.31)	(3.33)		(3.43)	(3.45)
TURN	0.029*	0.029*	ΔTURN	0.001	0.001	ΔTURN	0.000	0.000
	(1.68)	(1.68)		(0.29)	(0.29)		(0.33)	(0.32)
INTAN_F	-0.340	-0.338	ΔINTAN_F	0.970***	0.967***	ΔINTAN_F	0.386***	0.384***
	(-1.02)	(-1.01)		(3.54)	(3.53)		(3.38)	(3.37)
RD_F	2.242***	2.242***	ΔRD_F	1.954***	1.971***	ΔRD_F	0.891***	0.899***
	(2.96)	(2.97)		(2.72)	(2.74)		(3.01)	(3.03)
MOM	-0.009***	-0.009***	MOM	0.004***	0.004***	MOM	0.002***	0.002***
	(-13.03)	(-13.33)		(5.99)	(5.96)		(6.21)	(6.18)
Industry FE	Yes	Yes	Industry FE	Yes	Yes	Industry FE	Yes	Yes
Year FE	Yes	Yes	Year FE	Yes	Yes	Year FE	Yes	Yes
Obs	32003	32003	Obs	29548	29548	Obs	29548	29548
R-square	0.637	0.638	R-square	0.053	0.053	R-square	0.001	0.011