ETFs and Information Transfer Across Firms

Sanjeev Bhojraj
Cornell University
sb235@cornell.edu

Partha Mohanram
University of Toronto
partha.mohanram@rotman.utoronto.ca

Suning Zhang
Iowa State University
suningz@iastate.edu

Abstract: This paper examines the role that exchange-traded funds (ETFs) play in the transfer and dissemination of information across firms, using earnings announcements as a source of information. We find that the anomalous over-extrapolation of intra-industry information documented in prior research weakens with ETF ownership of firms and becomes insignificant with sector ETF ownership. At the ETF level, we find that sector ETFs have improved information transfer around earnings announcements, with greater followers' reaction to the leaders' announcements and lower subsequent reversals. In contrast, non-sector ETFs appear to drive the reversal effect observed in the overall sample. We also provide evidence that the previously documented reduction in ERCs for firms with high ETF ownership is attributable to more efficient information transfer rather than investor inattention. Finally, we show that ETF ownership at the firm level is associated with lower levels of post-earnings announcement drift, but only when the ETF ownership is sector based. In general, we find that sector ETFs are effective at transmitting factor (industry) information impounded in earnings news but general broad market ETFs are not very useful (and potentially detrimental). Our results highlighting the role of sector ETFs help reconcile conflicting results in the prior research as to whether ETFs foster greater market efficiency.

Keywords: ETF, Information transfer, Post earnings announcement drift, sector ETF

JEL Classification: G12, G14, M41, D53

We would like to thank seminar participants at the 6th Annual ABFER conference (Singapore), I.I.M Bangalore, George Mason University, Iowa State university, London Business School Accounting Symposium 2018, Purdue University, University of Illinois at Chicago, University of Michigan, University of Rochester, University of Waterloo and University of Windsor for their helpful comments. All errors are our own.

ETFs and Information Transfer Across Firms

Abstract: This paper examines the role that exchange-traded funds (ETFs) play in the transfer and dissemination of information across firms, using earnings announcements as a source of information. We find that the anomalous over-extrapolation of intra-industry information documented in prior research weakens with ETF ownership of firms and becomes insignificant with sector ETF ownership. At the ETF level, we find that sector ETFs have improved information transfer around earnings announcements, with greater followers' reaction to the leaders' announcements and lower subsequent reversals. In contrast, non-sector ETFs appear to drive the reversal effect observed in the overall sample. We also provide evidence that the previously documented reduction in ERCs for firms with high ETF ownership is attributable to more efficient information transfer rather than investor inattention. Finally, we show that ETF ownership at the firm level is associated with lower levels of post-earnings announcement drift, but only when the ETF ownership is sector based. In general, we find that sector ETFs are effective at transmitting factor (industry) information impounded in earnings news but general broad market ETFs are not very useful (and potentially detrimental). Our results highlighting the role of sector ETFs help reconcile conflicting results in the prior research as to whether ETFs foster greater market efficiency.

Keywords: ETF, Information transfer, Post earnings announcement drift, sector ETF

JEL Classification: G12, G14, M41, D53

1. Introduction

The last decade has seen a significant shift in the asset management landscape with the growth of exchange-traded funds (ETFs). As of the end of 2015, the ETF industry had assets under management (AUM) of roughly $2.9 trillion of which nearly $2 trillion was in the US (ICI Factbook 2016). In the 10 years ending 2015, ETFs have seen inflows of nearly $1.6 trillion. ETFs play a significant role in financial markets, particularly equity markets, constituting roughly 30 percent (23 percent) of US market trading by value (by volume).[1] Given the size of the AUM and the proportion of trading volume they represent, understanding the relation between ETFs and their constituents is important to gain insight into the benefits and costs of these instruments. This paper examines the role that ETFs play in the transfer and dissemination of information between constituent firms.

Ex-ante, the impact of ETFs on information transfer between its constituents is unclear. Information is comprised in varying degrees, of idiosyncratic components, industry level information and information about the economy or market level information. One of the biggest benefits of ETFs is that they allow for trading a large number of stocks in a cost-efficient manner. This feature of ETFs can allow commonality in information (like industry information or market level information) to quickly percolate through to all the constituents of the ETF. Trading using ETFs can be a more efficient mechanism to capture industry level information or market level information incorporated in any information source. This would suggest that ETFs would make stocks prices of its constituents more efficient. On the other hand, ETFs could be limited in their ability to transfer information effectively across firms if they are simply used as passive investment vehicles rather than as a means to efficiently trade on information. Further, even if they are used

[1] https://www.ft.com/content/6dabad28-e19c-11e6-9645-c9357a75844a

as vehicles for trading on factor information, they are limited in that they are baskets created based on fixed rules (e.g., market capitalization weighted). ETF trading thus could cause information to be impounded in the proportion determined by the rules rather than based on the value of the information to each constituent. In addition, ETF trading driven by earnings news that has no relevance to constituents other than the announcing firm can cause a mispricing in the other firms, which can persist, unless another group of investors (active managers or arbitrageurs) trade in the individual constituents to correct any mispricing.

Prior research examining the impact of ETF ownership and trading on the processing of information has found mixed evidence. Many papers paint a negative picture of ETFs by showing that ETF ownership and trading leads to increased return co-movement (Leippold, Su and Zeigler 2016, Da and Shive 2018), increased volatility and bid-ask spreads (Ben-David, Franzoni and Moussawi 2017) and reduced earnings response coefficients and analyst following (Israeli, Lee and Sridharan 2017). Conversely, Glosten, Nallareddy and Zou (2017) find that ETF membership improves the contemporaneous price-earnings relationship, especially among firms in poor information environments.

An important conditioning variable that the prior literature largely ignores is the nature of the ETFs, viewing them as a homogenous group. In reality ETFs can be broadly partitioned into at least two distinctly different groups. The first group is market level ETFs such as SPY (S&P 500 SPDR Fund) and VOO (Vanguard S&P 500 Index fund). The second group is sector or industry level ETFs such as XLK (Technology Select Sector SPDR Fund) and XLY (Consumer Discretionary Select Sector SPDR Fund). Subrahmanyam (1991) and Cong and Xu (2017) suggest that creating composite security designs that deviate from market weights allows for better factor investing. These securities could result in greater efficiency compared to market weighted ETFs, as they allow factor investors to better express their information advantage.

In this paper, we investigate the role of ETFs in facilitating the dissemination of relevant information contained in the earnings of a firm to other constituent firms. We also examine if the type of ETF (sector and non-sector ETFs) influences the role that ETFs play in facilitating earnings related information flow. Our research design focuses on the stock returns to firms around earnings announcements – both their own, as well as that of peer firms. We carry out our analysis at two levels. Firstly, at the firm level, we look at a broad sample of earnings announcements conditioned on the presence of and nature of ETF ownership. Secondly, at the ETF level, we identify the five largest holdings within each ETF and examine the returns of the first firm to announce earnings (the leader) and the other four firms (followers).

Our firm level analysis follows Thomas and Zhang (2008) who examine the intra-industry transfer of information around earnings announcements. They find that markets overestimate the industry level information in early announcers' earnings for late announcers' earnings and correct this overestimation when late announcers disclose their earnings. Using their research design, we examine whether the introduction of ETFs (particularly sector ETFs) has affected this overreaction by either exacerbating it or mitigating it. We begin by corroborating their finding using the universe of firms with available data in the 1985 to 2015 period. When we partition the sample into a pre-ETF period (1985-2001) when ETFs did not have a significant presence and a post-ETF period (2002-2015), we find a significant weakening of the intra-industry overestimation in the post-ETF period. We find that the results are primarily driven by sector ETFs. To mitigate the potential confounding impact of other events during the post-ETF period (e.g., Sarbanes-Oxley), we carry a narrow window analysis in the four quarters before and after a firm starts being owned by ETFs. We find that ownership by sector ETFs moderate the overestimation of intra-industry information, while there is no discernible effect for broad-based ETFs.

To explore this question further, we examine the relationship between earnings announcements of the firms that are owned by a given ETF. We identify the five largest holdings within each ETF for a large sample of ETFs from 2002-2015 and examine the stock returns of the first firm to announce earnings (the leader) and the four following firms (followers). We find that, on average, follower returns are positively associated with leader returns around the leader's earnings announcement. Further, this effect is much stronger for sector ETFs, especially when the level of ETF trading is high. We next test whether the ETF associated market reaction experienced by follower firms around the leader's earnings announcement is efficient or not by examining the returns in between the leader's earnings announcement and the follower's own earnings announcement. We find that followers in non-sector ETFs experience a larger reversal as compared with followers in sector ETFs. This suggests that the follower constituents of non-sector ETFs are more likely to experience an overreaction around the leader's earning announcement. Sector ETF followers experience a more muted reversal suggesting a more efficient price response around the leader's earnings announcement. Partitioning the sample based on the underlying ETF trading volume shows evidence consistent with the reversal being driven by high-volume ETFs.

We corroborate these findings by carrying out a path analysis that separates the overall correlation between the constituents into a direct correlation between the constituents and an indirect effect that flows through ETFs. The path analysis shows that ETFs play a significant role in facilitating the flow of information between their constituents. The results are particularly strong for sector ETFs, which show that the ETF channel contributes to roughly 61 percent of the overall correlation as compared with only 16 percent for non-sector ETFs.

Having documented that ETFs, specifically sector ETFs, can facilitate the transfer of information among individual constituents, we turn our attention to whether these information transfers account for the reduction in earnings response coefficients (ERCs) among high ETF ownership firms documented in the prior literature (Israeli, Lee and Sridharan 2017). We examine whether the decline in ERCs are a function of the ordering of earnings announcements and whether the decline varies between the two subgroups. We find that ERCs for follower firms decline as the fiscal quarter progresses. Further, this pattern of declining ERCs is only observed when the firms are ETF constituents. These results paint a different picture of the previously documented results on lower ERCs by attributing it to more efficient information transfer instead of less firm-specific investor attention, as interpreted by Israeli, Lee and Sridharan (2017).

To further distinguish between these two explanations, we examine the impact of ETF membership on the post-earnings announcement drift (PEAD), a well-established anomaly that shows that markets are slow to respond to information in earnings announcements. If the drop in ERCs is because of ETFs facilitating a more efficient information transfer, it should manifest in a lower PEAD as more information has already been impounded into price before the earnings announcement. If, however, the lower ERCs are driven by investor inattention, PEAD is likely to be unaffected or exacerbated by ETF ownership. We find that ETF ownership mitigates the drift – i.e. financial information is impounded into prices earlier for firms with greater ETF ownership. Crucially, this effect is driven entirely by sector ETF ownership. For non-sector ETF ownership, we see limited impact of ETF ownership on the drift. This supports our interpretation that ETF ownership, especially sectoral ETF ownership, facilitates information transfer across firms.

Our results hence bridge the seemingly inconsistent results in the prior academic literature regarding the impact of ETFs on market efficiency. By separating ETFs into sector ETFs and non-sector ETFs, we can compare ETFs that are better designed to impound industry level information in earnings to other kinds of ETFs. Non-sector ETFs can cause ETF constituents react to information that may not be relevant, causing anomalous return co-movement and future reversals, leading to increased return volatility. While sector ETFs also increase return co-movement, this is driven by the relevance of common information, leading to earlier impounding of news pertinent to future earnings, which in turn reduces earnings drift.

Our paper also contributes to the literature on intra-information industry transfer (Freeman and Tse 1992, Thomas and Zhang 2008). Our results suggest that the anomalous over-extrapolation of intra-industry information is reduced among firms with ETF ownership, and in fact becomes insignificant for firms with sector ETF ownership.

The rest of the paper is organized as follows. Section 2 provides institutional details regarding ETFs and discusses prior research that motivates our research questions. Section 3 describes the research design. Section 4 discusses the results from the firm level analysis, while section 5 discusses the ETF level analysis. Section 6 analyzes the potential impact of ETFs on the post-earnings announcement drift. Section 7 concludes the paper.

2. Institutional background and prior research on ETFs

2.1 The Emergence of ETFs

Exchange-traded funds (ETFs) are investment companies classified as open-ended companies or unit investment trusts (UITs). The first U.S ETF began trading in 1993 (SPY, S&P 500 SPDR) but they became very popular after 2000. ETFs assets in 1993 were 464 million dollars and grew to about 33 billion dollars by the end of 1999. ETF assets experienced significant growth from 2000 to 2011 rising to about one trillion dollars. In the four years from 2012 to 2015, these assets experienced phenomenal growth by doubling in size to two trillion dollars. As of the end of 2015 there were 1500 ETFs, up from about 80 in 2000. ETFs historically have tracked indices but more recently (starting in 2008), ETFs also include actively managed vehicles.[2]

An ETF is created by a sponsor who chooses the investment objective and benchmark for the ETF. This could be a market capitalization weighted index or other indices created using alternative techniques like equal weighting or factors such as value, growth etc. Index based ETFs could perfectly mimic the underlying index or choose a representative sample of stocks. For example, SPDR S&P 500 ETF Trust (SPY), which is the largest ETF, tracks the S&P 500 which is a float weighted index. Other ETFs could use alternative methods – e.g. First Dow Jones Internet Index Fund (FDN) tracks the Dow Jones Internet Index which is float and volume weighted, while the PowerShares Value with Momentum ETF is a factor-based ETF.

ETFs are unique investing vehicles that share some similarities and differences from open-ended mutual funds. They are similar to other open-ended funds because ETFs own the underlying assets (stocks, bonds, commodities, futures, foreign currency, etc.) and divide ownership of those assets into shares. However, unlike open-ended funds which are priced at the end of trading day, ETFs are derivative instruments that are traded on stock exchanges and priced throughout the day Investors can purchase ETF shares on margin, short sell shares, or hold for the long term.

The ability to trade easily, overlaid with low fees and diversification, results in ETFs being very popular and having high trading volume. ETFs on average represent roughly 30 percent of daily market volume. In 2016, the top twelve most traded securities were all ETFs, ahead of the

[2] While ETFs cover a wide spectrum of asset classes, they are predominantly equity focused. Of the approximately two trillion dollars in ETF assets as of the end of 2015, equity ETFs accounted for approximately 82 percent, bond and hybrid ETFs accounted for about 16 percent, and commodities for the remaining 2 percent. While most ETFs are passive vehicles that track an index active ETFs are gaining in popularity, though these are still tiny with an AUM of approximately $27 billion as of the end of 2015 (2016 ICI Factbook).

most traded individual security (AAPL). ETFs can trade throughout the day because of the transparency of their holdings. Unlike mutual funds that provide quarterly disclosure on holdings, ETFs provide holdings information daily in disclosures referred to as creation baskets. As a result, unlike the mutual funds for which NAV is known only at the end of the day, the NAV for ETFs is available at any time during the trading day.

2.2 Primary and Secondary ETF Trading

ETFs are subject to two types of trading: primary market trading and secondary market trading. Trading in the primary market involves the process of creation and redemption of ETFs. The sponsor manages the process of creating and redeeming ETFs through a group of intermediary financial institutions called authorized participants (AP). If the demand for ETFs exceeds the available shares, an AP can buy the underlying constituent portfolio and deposit it with the sponsor in exchange for shares in the ETF (ETF creation). Similarly, if the supply of ETFs in the market exceeds the demand, an AP can buy ETFs in the open market and give them to the sponsor in exchange for shares in the underlying constituents (ETF redemption). It is worth noting that on average, about 10 percent of the daily ETF volume occurs in the primary market.

Trading in the secondary market is the dominant form of trading in ETFs accounting for about 90 percent of ETF volume. Secondary ETF trading occurs without the creation of new ETF shares or the redemption of existing ETF shares. Secondary ETF trading can affect that price of the ETF and by extension the price of the underlying securities. For example, if buyers exceed sellers, it puts upward price pressure on the ETF and causes a spread between the price of the ETF and the NAV based on the underlying stocks. APs and other market participants ensure that the spread is quickly corrected by buying the underlying the basket while simultaneously selling (or short-selling) the ETF. Trading in both the primary and the secondary market will have an impact on the pricing of the underlying constituents. To create more ETF shares or redeem existing shares, the APs have to buy or sell the underlying constituents causing the price of the underlying constituents to rise or fall. Similarly, to arbitrage price differences in the secondary market, arbitrageurs will have to buy or sell/short-sell the ETF and sell/short-sell or buy the underlying constituents again causing the price of the underlying constituents to move based on ETF activity.

2.3 How ETF Trading can influence the efficient pricing of the underlying securities

One of the biggest benefits of ETFs is that they allow for trading a large number of stocks in a cost-efficient manner. Subrahmanyam (1991) and Cong and Xu (2017) argue that ETFs are particularly useful to factor-informed traders as they can easily trade on their information advantage.[3] For example, prior to the introduction of market-wide ETFs (e.g., SPY), investors who possessed information that could affect the overall market had a high barrier to trading on their information as they would have to trade (long or short) several hundred individual stocks to express their information. ETFs make this process easier by reducing transaction costs as well as the speed with which the transactions are executed. Given that most of the ETF trading occurs in the secondary market, ETFs reduce the constraints that can arise due to limited liquidity in some of the underlying constituents. ETFs also make it easier to express negative information through shorting because it significantly eases the locate process (one locate instead of several)[4] and the cost of shorting (Huang, O'Hara and Zhong, 2018).

By lowering the bar for trading by factor-informed investors, ETFs allow commonality in information (like industry or market level information) to quickly percolate through to all the

[3] Examples of factor information could include industry level information and information about the economy or market level information.

[4] Locate refers to the process through which institutional investors obtain the shares for short selling.

constituents of the ETF. Trading using ETFs can be a more efficient mechanism to capture industry level information or market level information incorporated in any information source. Consistent with this, Glosten, Nallareddy and Zou (2017) find that ETFs lead to more timely impounding of systematic information, especially for firms in weak information environments.

ETFs could also contribute negatively to efficient pricing. As ETFs are baskets created based on fixed rules (e.g., market capitalization weighted), they could be limited in their ability to transfer information effectively across firms. When factor-informed investors use ETFs, they force information to be impounded based on the proportion set by the rules rather than based on the value of the information to each of the constituents. ETF trading can cause anomalous movement in the stock price of underlying securities, even if the information is irrelevant for a given security. This can result in the constituents being potentially mispriced unless another group of investors (active managers or arbitrageurs) trading in the individual constituents quickly correct this mispricing. Cong and Xu (2017) suggest that that increased ETF ownership may dis-incentivize traders from acquiring information of individual stocks, leading to fewer firm-specific informed traders and potentially greater pricing inefficiency. Consistent with this, Israeli, Lee and Sridharan (2017) find that greater ETF ownership results in lower analyst following.

2.4 Broad-based ETFs vs Sector ETFs

While prior studies have viewed ETFs as a homogenous group, in reality ETFs can be broadly partitioned into at least two distinctly different groups. The first group is market level ETFs such as SPY (S&P 500 SPDR ETF) and VOO (Vanguard S&P 500 ETF). The second group is sector or industry level ETFs such as XLK (Technology Select Sector SPDR ETF) and XLY (Consumer Discretionary Select Sector SPDR ETF). Subrahmanyam (1991) and Cong and Xu (2017) suggest that creating composite security designs that deviate from market weights or expressing factor weights that are different from market weights allows for better factor investing. At the end of 2016, sector ETFs were the top five highest turnover ETFs and accounted for eight of the top 10. These securities could result in greater efficiency compared to market weighted ETFs, as they are better designed to capture industry information.

Sector ETFs are likely to contribute to more efficient transfer of information as their underlying firms come from a more homogenous group. Thus, when one firm releases news, the odds that the news is relevant for other firms in the ETF is higher. Conversely, non-sector ETFs consist of more heterogeneous constituents, which makes it more likely that the information released by one firm is irrelevant for other firms. In our analysis, we will consider the important conditioning role of the nature of the ETF to attempt to reconcile the seemingly contradictory findings in prior research.

3. Research Design

3.1 Analysis of Returns around Earnings Announcement

We examine the role that ETFs play in information transfer by focusing on earnings announcements as a source of information. Different constituent firms within an ETF release earnings information at different points of time within a given quarter. This provides us with an opportunity to study whether other ETF constituents react when a firm within an ETF releases earnings. This also allows us to examine whether the initial reaction was efficient or not, by examining subsequent returns.

We carry out our analysis at two levels – the firm level and the ETF level. Firstly, at the firm level, we look at a broad sample of firm-level earnings announcements. We then examine a well-established result in the information transfer literature: the anomalous over-extrapolation of industry information as documented by Thomas and Zhang (2008). We test whether ETF ownership and the nature of ETF ownership (sector vs non-sector) is associated with a change in the pattern of over-extrapolation. Secondly, we also carry out analysis at the ETF level by identifying the five largest holdings within each ETF and examining the returns of the first firm to announce earnings (the leader) and the other four firms (followers). We examine information transfer by studying how followers react to the earnings announcement of leaders and whether this reaction varies by the nature of the ETF. We also examine whether this reaction was efficient or anomalous by studying the returns in the period between the leader's announcement and the followers own earnings announcement.

3.2 ETF Sample Construction

Our sample selection procedure is outlined in Table 1. Panel A describes how we arrive at the sample of ETFs for our analysis. As discussed in section 2, an ETF is a type of fund that owns the underlying assets (stocks, bonds, commodities, futures, foreign currency, etc.) and divides ownership of those assets into shares. For our study, we focus on ETFs with underlying assets in shares of stocks (i.e., equity ETFs). First, we use CRSP to identify ETFs traded on major US exchanges (CRSP historical code of 73). ETFs are required to disclose their portfolio holdings at the end of each quarter on SEC forms N-CSR and N-Q. We hence merge the names of the ETFs with Thomson-Reuters Mutual Fund Holding (S12) database to construct ETF holdings for each stock at the end of each quarter. This process yields 487 ETFs in the period from 2002 to 2015,[5] which is similar to the number of ETFs from the literature (Israeli, Lee and Sridharan 2017, Glosten, Nallareddy and Zou 2017).

[5] We start our ETF level analysis from 2002, since ETF ownership was low before 2002. In addition, SEC proposed to require funds to file their complete portfolio holdings schedules with the Commission on a quarterly basis, rather than semi-annually in 2002.

We identify sector ETFs by analyzing the title of the ETF, as sector ETFs typically specific sectors and industries. NYSE, NASDAQ and some popular ETF websites, such as ETF.com and ETFdb.com give a comprehensive list of the names of sector ETFs. We first rely on these names to code our ETFs and also conduct a final check by reading the name of each ETF to prevent miscoding. Of these 487 ETFs, 214 were sector ETFs while 273 were broader non-sector ETFs. For comparison, Huang, O'Hara and Zhong (2018) identify 217 industry ETFs before narrowing their sample down based on additional screens.

Panel B displays the sample distribution across time. The sample has noticeably fewer observations in the early years, with 106 distinct ETFs in 2002. This increases sharply to 413 distinct ETFs by 2007 and 473 distinct ETFs in 2008, declining slightly after that. Overall, there is little evidence of time clustering in our sample. Panel C presents the distribution of the different sectors that the sector ETFs in our sample focus on. The 214 sector ETFs can be classified into 32 distinct sectors or subsectors. While many sectors have only a handful of ETFs, a few sectors have a large number of competing ETFs - e.g. Consumer Products (21), Energy (10), Financial Services (15), Healthcare (29), Real Estate (20) and Technology (18).

4. Firm Level Analyses

Thomas and Zhang (2008), henceforth TZ, document that investors overestimate the intra-industry implications of early announcers' earnings for late announcers' earnings and that this overestimation is corrected when late announcers disclose their earnings. If ETFs have led to more efficient information transfer, we should see a weakening in this overestimation as ETF ownership has increased, especially for sector ETFs. TZ define peer firms (early announcers) for any particular firm as the firms in the same industry that report earnings at least 5 days preceding that firm's earnings announcement date. Variable *ARET* represents the size-adjusted excess return for the 3 day window around the announcer firm's own earnings announcement. Since there is typically more than one peer firm for each firm, variable *RESP* is the average excess return of the firm around its peer firms' earning announcement dates. They run the following specification that controls for other determinants of announcement period returns.[6]

$$ARET = \alpha + \beta_1 {*} RESP + \beta_2 {*} ACC + \beta_3 {*} RET6 + \beta_4 {*} ERLYPRARET + \beta_5 {*} ARET1 + \beta_6 {*} ARET4 + \beta_7 {*} INST + \beta_8 {*} SIZE + \beta_9 {*} LOGBM + \varepsilon \qquad (1)$$

TZ document a significant negative relationship between *ARET* and *RESP* confirming that investors overestimate the intra-industry implications of early announcers' earnings for late announcers' earnings. We begin with a replication of TZ with one modification – we use the Fama and French (1997) industry classification instead of four-digit SIC code, as it potentially better reflects the composition of firms within sector ETFs (e.g. Technology ETF, Telecommunications ETF, Biotech ETF).[7] Firms' quarterly earnings announcement dates are from quarterly Compustat files. Both *ARET* and *RESP* measure excess returns accumulated over 3 trading days starting from the trading day before earnings announcement date, computed as raw returns minus the returns from the same NYSE/AMEX/NASDAQ size decile firms over the same event window.

Panel A of Table 2 presents the results. The first column presents the results for the entire sample (1985-2015). Consistent with TZ, we find a significant negative relationship between *ARET* and *RESP*, with *RESP* having a coefficient of -0.1950 (t-statistic -10.78). We next partition our sample into a pre-ETF period (1995-2001) where ETF ownership of stocks was virtually

[6] The control variables are accruals (*ACC*), buy-and-hold returns for the six months prior to announcement (*RET6*), average of early peer's three day earnings announcement excess returns (*ERLYPRARET*),one quarter and four quarter lagged announcement returns (*ARET1*, *ARET4*), institutional ownership percentage (*INST*), log of market capitalization (*SIZE*) and log of the B/M ratio (*LOGBM*).

[7] Results are similar if we use 4-digit SIC codes to identify peer firms.

nonexistent and the ETF period (2002-2015). We find that the reversal in returns is much stronger in the pre-ETF period (-0.2315, t-stat -10.01) than the ETF period (-0.1488, t-statistic -11.78). The final regression shows that the difference between these coefficients is highly significant (0.0807, t-stat 3.10). This provides prima-facie evidence that the overreaction to industry information has weakened as ETFs have become more prevalent. However, this does not provide a direct link, as this period was also associated with numerous other changes in various aspects of the capital markets such as market microstructure (decimalization), regulation (Sarbanes-Oxley), availability of peer information (EDGAR, XBRL) and functioning of analysts (global settlement).

To better link the change in intra-industry information transfer to ETF ownership, we consider the impact of ETF ownership at the firm-level. We only consider firms that had ETF ownership at some point in our sample. We measure ETF ownership by summing up the total ownership by all ETFs to construct the percentage of shares owned by ETFs at the end of each firm quarter.[8] After identifying the first instance of ETF ownership, we study the narrow window around the initiation of ETF ownership – the four quarters before and after ETF ownership. This ensures that the firm acts as its own control, and we can have greater confidence in ascribing any effects we identify as being associated with ETF ownership. Further, pre and post ETF are not fixed but vary by firm based on when ETFs started to own shares in a given firm. Finally, the short window increases the confidence that any differential effect we find is more likely to be associated with ETF ownership rather than other confounding events.

For each firm, we create the following three indicator variables. *ETF* is an indicator variable that equals 1 if ETFs own stock of the given firm in the given quarter and 0 otherwise. By construction, each firm will appear eight times in the regression, four times with *ETF*=0 and four

[8] Since ETFs only report their holdings at the end of March, June, September, and December, we require firms with fiscal quarter ending aligned with calendar quarter.

times with *ETF*=1. *ONLYBROAD* is an indicator variable that equals 1 if the ETFs that own stock in the given firm are entirely non-sector and 0 otherwise. *SECTOR* is an indicator variable that equals 1 if sector ETFs own stock of the given firm in the given quarter and 0 otherwise. In our regressions, we interact *RESP* first with *ETF*, and then with *ONLYBROAD* and *SECTOR*. If ETF ownership mitigates the over-extrapolation that TZ document, we expect the interactions to have a positive coefficient, with a stronger effect for sector ETFs.

The results are presented in Panel B of Table 2. Column 1 presents the regression using the interaction of *RESP* with *ETF*. The main effect on *RESP* represents the reversal in the immediately pre-ETF period, while the coefficient on *RESP*ETF* represents the change in the reversal in the immediately post-ETF period. *RESP* has a significant negative coefficient (-0.2434, t-stat -9.13) confirming the baseline TZ result. The interaction of *RESP*ETF* has an insignificant positive coefficient (0.0491, t-stat 0.82) suggesting that ETF ownership does not significantly mitigate intra-industry overestimation. Column 2 presents the regression using the interaction of *RESP* with *ONLYBROAD* and *SECTOR*. Consistent with our expectations, we find greater mitigation when the ETF ownership includes sector ETF ownership. The interaction of *RESP*SECTOR* has a significant positive coefficient (0.1074, t-stat 3.07), while the interaction of *RESP*ONLYBROAD* is insignificant (0.0136, t-stat 0.18). Hence, we find a significant weakening of the intra-industry over-extrapolation when sector ETFs own shares in firms, but fail to find an effect for broader ETFs. These results consistent with sector ETFs being more effective as a mechanism to transfer information between its constituents.

5. ETF Level Analyses

To better understand the dynamics of information propagation within ETFs, we analyze the returns around earnings announcement for the firms that are owned by a given ETF. ETFs often own a large number of stocks, including small capitalization stocks that are illiquid and thinly traded. Given our interest in the pricing of earnings related information, we focus our attention on the five largest holdings in a given ETF based on dollar value weights.[9] The first firm to release quarterly information is referred to as the leader, while the other four are referred to as followers.[10]

5.1 Sample and Variable Definitions

The sample of 487 ETFs correspond to 16,707 distinct ETF-quarters from 2002 to 2015.[11] We begin by identifying the top 5 holdings of each ETF by dollar value weight of each constituent at the end of each quarter, and pair the leader with four followers. We only keep firms with fiscal ending aligned with calendar ending (i.e. quarters ending in March, June, September and December) and no missing underlying ETF trading volume information. We delete observations where the follower's announcement was within two days of the leader, as well as observations with missing earnings announcement date and timestamps on IBES. We adjust the earnings announcement date of the firms that announce earnings after market closes to next trading day, following the recent study by Beaver, McNichols and Wang (2017). We also delete observations where we are unable to obtain ETF trading volume. Our last data step is to remove duplicate pairs

---

[9]In a recent working paper Pan and Zeng (2017) find mispricing that occurs because of a liquidity mismatch between liquid bond ETFs and illiquid underlying bond instruments. In a related paper Bhattacharya and O'Hara (2017) examine the informational efficiency of underlying markets when the constituents underlying ETFs are illiquid. Focusing on the top five ETF constituents helps avoid this issue. We also limit our analysis to the two largest holdings, i.e. with one leader and one follower, with similar findings.

[10] As an alternative approach we also identify leaders as the largest market capitalization firm from among the five stocks and followers as subsequent announcers. The results are similar to those discussed using the main sample.

[11] We start our analysis from 2002 as that is when quarterly reporting data on ETF holding starts.

– the fact that two firms in a pair might be owned by multiple ETFs. We apply two-step screens to delete repeated leader-follower pairs in any given quarter. For repeated pairs, we keep the pair that is classified as high-volume and/or the pair that is from a sector ETF. The final sample has 31,922 leader-follower pairs from 9,851 distinct ETF-quarters.

We define *LRET* as the size-adjusted returns accumulated over 3 trading days starting from the day before leader's announcement date. For each follower, we compute market responses over 2 windows. $FRET_{ANNC}$ measures follower's response to leader earnings announcement, computed as the size-adjusted returns accumulated over 3 trading days starting from the leader's announcement date. $FRET_{BETW}$ measures follower's size-adjusted stock returns over the period between the earnings release of the leader and the earnings release of the follower firms[12]

5.2 Analysis of returns around leader's earnings announcement

We begin our analysis by studying the investor response to the event window around the leader's earnings announcement. The results are presented in Table 3. In all our regressions, the t-statistics control for two-way clustering at the leader and leader's earnings announcement date level, because the same leader may be paired with multiple followers.

Panel A of Table 3 examines the relationship between the leader's stock returns (*LRET*) and the follower firms' stock returns ($FRET_{ANNC}$). For all firms, the coefficient on *LRET* is 0.049 (t-stat 8.51), which increases to 0.085 (t-stat 7.95) for sector ETF pairs and decreases to 0.023 (t-stat 3.69) for non-sector ETFs. These results suggest that follower firm's stock returns are associated with the leader firm's stock returns and that this relationship is stronger for sector ETFs.

---

[12] Size-adjusted returns are the daily stock return in excess of the return on a value weighted portfolio of firms having similar market values. The size portfolios are formed by CRSP and are based on size deciles of NYSE and AMEX firms. Membership in a particular portfolio is determined using the market value of equity at the beginning of the calendar year.

If ETFs do play a role in the response of follower firms to the leader's earnings announcement, we expect to see follower-leader pairs with high ETF trading volume to have a greater impact as compared with the pairs with low ETF trading volume. To better link our findings with ETF trading, we carry out cross-sectional analysis during the ETF period by partitioning ETFs into above and below median groups based on their trading volume over the 3 trading days around leader's earnings announcement date [-1,1]. We define a leader-follower pair to be high-volume if the average trading volume of the underlying ETF over the 3 trading days around leader's earnings announcement date [-1,1] is above the median ETF average daily trading volume in the quarter before earnings season starts. The results are provided in panel B of Table 3. The coefficient on *LRET* is similar for high volume and low volume ETFs suggesting that the response of non-sector followers to the leader's announcement is not determined by ETF trading volume. However, Columns 2 and 3 of the panel B of Table 3 show that the sector ETF pair results seen in panel A of Table 3 are primarily driven by high volume ETFs. The coefficient on the interaction term *LRET*SEC* is large and highly significant for high volume ETFs (0.103, t-stat 5.46) and insignificant for low volume ETFs (0.017, t-stat 1.21). Column 4 repeats the analysis and shows that the difference between low and high volume groups is significant (0.086, t-stat 3.68).

Overall, the results in Table 3 suggest that while firms within an ETF react significantly to the early announcing firm's earnings announcement, this effect is particularly strong for sector ETF firms, and when ETF trading volumes are higher. These results are also consistent with the increased return co-movement associated with ETFs documented in prior work (e.g., Da and Shive 2018, Israeli, Lee and Sridharan 2017), though our analysis to follow suggests a different explanation for this increased co-movement.

5.3 Analysis of returns between leader's and followers' earnings announcement

The results in Table 3 suggest that ETFs are associated with the extent to which follower firms react to the leader's earnings announcement. They do not however examine whether this reaction is "efficient". Indeed, the reaction shown in Table 3 could well be anomalous overreaction, akin to the RESP variable used in our firm-level analysis. To test this, we examine the followers' returns in the period between the leader's and the followers' earnings announcement. We measure the size-adjusted returns ($FRET_{BETW}$) for follower firms in the period between the leader's earnings announcement and the follower firm's own earnings announcement. We test the correlation between $FRET_{BETW}$ and $FRET_{ANNC}$. If we find a negative correlation, it suggests a reversal of the initial reaction.

The results are presented in panel A of Table 4. The first column presents the results for the entire sample. We find a significant negative coefficient on $FRET_{ANNC}$ (-0.071, t-stat -3.89). This suggests that the markets correct their initial reaction to the leader's earnings announcement. However, this reversal is primarily driven by non-sector ETFs with the coefficient on $FRET_{ANNC}$ of -0.109 (t-stat -3.54). For sector ETFs, the evidence of a reversal is much weaker, as the coefficient on $FRET_{ANNC}$ is lower in magnitude at -0.037 (t-stat -1.89). The difference in reversal between the two groups is statistically significant (t-stat 1.99). Sector ETFs, which are generally more focused on stocks with common factors, are likely to be more efficient in transmitting information to its constituents, as observed by the greater reaction around the leader's earnings announcement return and the weaker reversal in subsequent days. Non-sector ETFs on the other hand seem to be associated with an overreaction in follower returns, consistent with trading in these ETFs potentially causing a mispricing of the follower stocks.

If ETFs do play a role in the return reversals of follower firms subsequent to the leader's earnings announcement, we expect to see a stronger effect on follower-leader pairs with high ETF trading volume. The results are provided in panel B of Table 4. The coefficient on $FRET_{ANNC}$ is larger for high volume than low volume pairs suggesting that the response of non-sector followers to the leader's announcement is determined by ETF trading volume (column 1). More importantly, Columns 2 and 3 show that the sector ETF results in panel A of Table 4 are primarily driven by high volume ETFs. The coefficient on $FRET_{ANNC}*SEC$ is large and statistically significant for high volume ETFs (0.104, t-stat 1.83) and insignificant for low volume ETFs (0.045, t-stat 0.99) albeit a test of the difference between these two coefficients is not statistically significant. If the findings in panel A of Table 4 were simply related to intra-industry information transfer and not ETF membership, we would expect both groups to have the same response. This finding is consistent with sector ETFs having facilitated factor investing, resulting in smaller return reversal after leader's earnings announcement. This also corroborates our firm-level analysis, where we find that the over-extrapolation of intra-industry information is insignificant for sector ETFs.

Together these results are consistent with the general message that results on ETFs are contextual to the nature of information and the kinds of ETFs examined. Sector ETFs have improved the information environment by facilitating flow of common information while non-sector ETFs appear to associated with potential mispricing in the underlying securities.

5.4 Path Analysis of the mediation effects of ETFs

We attempt to isolate the role of ETFs in the flow of information between the constituent firms by carrying out path analysis. Using an approach similar to Bhattacharya et al. (2012), we decompose the correlation between the source variable (returns to the leader around its own earnings announcement, *LRET*) and the outcome variable (returns to the follower around the leader's earnings announcement, $FRET_{ANNC}$) into two paths, direct path and mediated path. [13] The co-movement between the leader and followers could come directly from the leader (direct path or direct correlation between leader and follower returns) or could be mediated by the existence of ETFs (indirect path through ETF returns). In Figure 1 we provide the diagram showing the two paths. We use a recursive approach where all paths flow in one direction i.e., from the leader to the follower either directly or through the mediating influence of the ETF.

The results from this analysis are provided in Table 5. Using the entire sample of ETFs, we find that they play a significant role in the relationship between leader and follower returns. Panel A of Table 5 provides information on the total effect and the breakup between the mediated effect and the direct effect. The total correlation is 0.097, which can be parsed out into a direct effect of 0.055 and a mediated effect of 0.042 (a product of the correlation between the leader return and ETF return (0.170) and the correlation between ETF return and follower return (0.245). Thus, about 40 percent of the total correlation between the returns is through ETFs while the rest is through a direct effect of leader on follower returns.

Parsing out the sample into sector and non-sector ETFs yields insights consistent with our prior tables. The total correlation between leader and follower returns is 0.149, much higher than the correlation for the overall sample. What is more interesting is that 61 percent of this correlation can be attributed to the mediating effect of ETFs. This is in contrast with the overall correlation for non-sector ETFs of only about 0.050, of which only 16 percent is affected through ETFs. These results are consistent with sector ETFs being more effective in transmitting information to its constituents as compared with non-sector ETFs.

[13] For a more detailed description of path analysis please see Asher (1983).

5.5 Falsification Tests

To bolster our results on the role of ETFs in the flow of information, we carry out falsification tests using a different sample period. We focus on the sample of leader and follower firms that are identified during the ETF sample period and go back in time to the pre-ETF sample period. Specifically, we consider the same firm-pairs in the 1985-1999 period, ensuring that they were not a part of an ETF at that time. We carry out an analysis similar to Tables 3 -4 for these firms in the pre-ETF sample period. If ETFs play a role in facilitating the dissemination of information, we should expect to see weaker links between the leader and follower returns in this period. The results are presented in Table 6.

Panel A of Table 6 replicates the analysis in Table 3. Consistent with the earlier results, we find that follower firm's returns do move when the leader announces earnings and that this effect is stronger for firms in the same sector. Panel B of Table 6 replicates the analysis in Table 4 and finds results that are markedly different. In the pre-ETF sample period (i.e. the falsification sample), we see no reversal in returns subsequent to the leader's earnings announcement, which is different from the ETF period returns. When parsing out the sample into sector and non-sector firms, we find that sector ETF follower firms experience significant post leader earning reversals. However, it is different from the results in Table 4 of the paper using the ETF sample wherein no significant reversal is observed. The reversal that is observed in the ETF sample period is much stronger for non-sector ETFs, while in the pre ETF period, the reversal is stronger for sector ETFs (i.e. firms in the same sector). These falsification results also corroborate our firm-level results as they show evidence of intra-industry overreaction as documented by Thomas and Zhang (2008) without the mitigating impact of ETF trading.

6. ETFs, Earnings Response Coefficients and the Post-Earnings Announcement Drift

6.1 Impact of ETFs on earnings response coefficients

The results thus far suggest that ETFs, especially sector ETFs, allow for the effective transmission of relevant information among ETF constituents. The firm-level tests suggest that ownership by sector ETFs reduce anomalous overreaction to industry peer firms' earnings information. The ETF level tests confirm this finding and show that there is a smaller reversal in the time between the leader's earnings announcement and the follower's earnings announcement for sector ETFs. The release of earnings information by any peer firm also releases correlated information, which can get impounded into stock prices through ETF trading. A corollary of this finding is that earnings response coefficients (ERC) should decline over the quarter. For early announcers, there is less common information that has already been released which makes the earnings release more informative. For late announcers, a portion of the information has been "pre-released" to the market which potentially makes the earnings release potentially less informative. Based on our earlier results that sector ETFs facilitate the flow of information between its constituents, we expect to find that the decline in ERCs should be stronger for firms that are constituents of these ETFs.

We test this conjecture using the following research design. We look at the universe of firm-quarters in the 1985-2015 period that have fiscal quarter ends in March, June, September and December. We consider the top five firms by market capitalization in each of the Fama and French (1997) 48 industry groupings. We rank these five firms based on their earnings release dates (RDQ) and create a rank variable called RRDQ that increases from zero for the first firm to four for the last firm. We run the following ERC regression, controlling for the determinants of announcement period returns, similar to our earlier TZ tests.[14] The sample consists of 24,472 firm-quarters for which complete data is available.

$$ARET = \alpha_0 + \beta_1{*}SURP + \beta_2{*}RRDQ + \beta_3{*}SURP{*}RRDQ + \beta_4{*}ACC + \beta_5{*}ARET1 + \beta_6{*}ARET4 + \beta_7{*}RET6 + \beta_8{*}INST + \beta_9{*}SIZE + \beta_{10}{*}LOGBM + \varepsilon \qquad (2)$$

In the above regression, the coefficient on earnings surprise (*SURP*) represents the earnings response coefficient for the first firm to release earnings information. The coefficient on the interaction (*SURP*RRDQ*) represents the trend in earnings response coefficients for later releasers.

The results are presented in Table 7. The first column presents the results for the entire sample. As expected, the coefficient on SURP is positive and significant (2.6634, t-stat -0.2363). The coefficient on *SURP*RRDQ* is negative and significant (-0.2363, t-stat -2.32), suggesting that ERCs decline for later releasers. This potentially represents transmission of industry specific information from early releasers to late releasers, given that we are analyzing firms within a given industry and are therefore capturing intra-industry information transfer.

To test whether the decline in ERCs is related to ETFs, we partition our sample into two subgroups. The first subgroup consists of observations where none of the five firms in the industry belong to any ETF. The second subgroup consists of observations where all five firms belong to an ETF. The next two columns present the ERC regressions for these two subgroups. For the no-ETF subsample, the coefficient on *SURP*RRDQ* is insignificantly different from zero (-0.1303, t-stat -1.29). For the all-ETF subsample, the coefficient on *SURP*RRDQ* is -0.491 (t-stat -3.80), suggesting a significant decline in ERCs as the quarter progresses.[15] The final two columns

[14] The control variables are accruals (*ACC*), buy-and-hold returns for the six months prior to announcement (*RET6*), one quarter and four quarter lagged announcement returns (*ARET1, ARET4*), institutional ownership percentage (*INST*), log of market capitalization (*SIZE*) and log of the B/M ratio (*LOGBM*).

[15] The coefficient on *SURP* is significantly higher for the all-ETF subsample as compared to the no-ETF sample. This can largely be attributed to the trends of increasing ERCs across time as documented by Beaver, McNichols and Wang (2017). The no-ETF sample has no observations after 2003 given that we focus on the five largest firms in an industry

consider two subgroups based on sector ETF ownership. We find an even stronger declining trend in ERCs when all five firms have sector ETF ownership, with the coefficient on *SURP*RRDQ* (-0.7772, t-stat -3.47).

To better illustrate these results, we present them graphically in Figure 2. For each group, the coefficient on *SURP* is the ERC for the first firm. For each subsequent firm, we infer the ERC by subtracting the value of the interaction (*SURP*RRDQ*) from the ERC of the previous firm. As the graph indicates, we see a sharp decline in ERCs when all five firms in an industry are part of an ETF, and especially when they are a part of a sector ETF. This is consistent with information transfer across ETF constituents prior to earnings release and hence provide an alternative explanation for the lower ERCs that Israeli, Lee and Sridharan (2017) document. While their explanation is based on investor inattention, our results are consistent with ETF trading providing a conduit for the impounding of relevant information when peer firms release earnings. Our final tests shed additional light on distinguishing between these two explanations.

6.2 Impact of ETFs on the post earnings announcement drift

The post-earnings announcement drift (PEAD) refers to the positive correlation between the returns in the period after earnings is announced and the earnings surprise (Bernard and Thomas, 1989, 1990). This is considered an anomaly because the return drift persists for a considerable period – as long as sixty days after earnings release. The most commonly accepted wisdom about this anomaly is that it represents the delayed processing of earnings information by capital market participants.

---

and the ubiquity of ETF ownership in the later years of our sample. Conversely, the all-ETF sample is concentrated in the latter years of our sample.

Our results thus far suggest that ETFs, especially sector ETFs, have a significant impact on the processing of earnings information by capital markets, by facilitating the impounding of correlated information released by peer firms. If more earnings information is impounded by the markets by the time of the earnings release, we should see a weaker PEAD. Alternatively, if markets are inattentive to firm-level information as Israeli, Lee and Sridharan (2017) conclude, we should see stronger PEAD in the presence of ETF ownership.

In our analysis, we analyze the impact of the extent of ETF ownership on PEAD. Panel A of Table 8 outlines how we construct the sample for the PEAD analysis. As discussed in section 3, the merging of equity ETFs with Thompson-Reuters Mutual Fund Holding (S12) yields 487 equity ETFs in the period from 2002 to 2015 and we are able to compute shares owned by each of these 487 ETFs at the end of each quarter. We then sum the total ownership by all ETFs to construct the percentage of shares owned by ETFs at the end of quarter for each firm quarter for all US incorporated firms that are traded on NYSE, NASDAQ, and AMEX. Consistent with our earlier analyses, we require firms that have fiscal quarter ends in March, June, September and December. After requiring non-missing values for firm characteristics that have found to be related to PEAD, we have 133,971 firm quarter observations over 2002 to 2015 period.

Earnings surprises (*SURP*) are calculated as I/B/E/S actual earnings per share minus the last mean analyst consensus forecast before the earnings-announcement dates, scaled by the stock price at the start of the quarter. We obtain our daily stock returns and daily stock prices from CRSP. To calculate the cumulated size-adjusted returns following earnings announcements (*POST60*), we require a firm to have a minimum of 40 days during the 60 trading days following the quarterly earnings announcements. We adjust the earnings announcement date of the firms that announce earnings after market closes to the next trading day, following the recent study by Beaver, McNichols and Wang (2017). The size portfolios are formed by CRSP and are based on size deciles of NYSE/NASDQ/AMEX firms. Membership in a particular portfolio is determined using the market value of equity at the beginning of the calendar year.

Panel B of Table 8 presents some summary statistics related to ETF ownership. The average firm in our sample has a mean 4.2% (median 3.4%) of shares outstanding owned by ETFs. The percentage of shares owned by sector ETF has a mean of 0.5% (median 0.1%). Figure 3 graphically presents the ownership by ETFs over time – for both ETF ownership in general and sector ETF ownership, one can observe a steady increase in ownership. Interestingly, we see a slight increase in the relative proportion of sector ETFs in recent years.

Panel C of Table 8 presents the average returns in the 60 days after earnings announcement (*POST60*) for deciles formed on the basis of the earnings surprise (*SURP*). The first column presents results for all firms. We see a monotonic increase in *POST60* across deciles of *SURP*, from -1.20% for the lowest SURP group to +3.14% for the highest *SURP* group, a spread of 4.34% that is significant at the 1% level, corroborating the evidence from the long-standing literature on PEAD (Bernard and Thomas, 1989, 1990). The next three columns consider sub-samples partitioned on the basis of ETF ownership. For the subset of firms that are owned by any ETF, the spread reduces to 4.23%, while for the subset with sector ETF ownership, the spread is only 3.22%, significantly less than the spread for all firms. For the subset of firms that are owned by broad ETFs, the spread worsens to 5.64%, much higher than the firms owned by sector ETFs. This provides preliminary evidence that sector ETF ownership is associated with lower drift, while broad ETF ownership is associated with higher drift.

We test the association between PEAD and ETF ownership in a multivariate regression setting, controlling for other known determinants of PEAD identified in prior research (Huang, Li and Wang 2015).[16] The regression specification is as follows.

$$POST60 = \alpha + \beta_1 {*}RSURP + \beta_2 {*}ETF + \beta_3 {*}RSURP{*}ETF + \beta_4 {*}SIZE + \beta_5 {*}BETA + \beta_6 {*}MTB + \beta_7 {*}PRERET$$
$$\beta_8 {*}RSURP{*}SIZE + \beta_9 {*}RSURP{*}BETA + \beta_{10} {*}RSURP{*}MTB + \beta_{11} {*}RSURP{*}PRERET + \varepsilon \quad (3)$$

To more easily interpret the coefficient on earnings surprise, we construct variable *RSURP*, where the 10th decile *RSURP* equals 1 and the 1st decile *RSURP* equals 0 (2nd decile RSURP=0.111, 3rd decile RSURP=0.222 etc). Thus, the coefficient for *RSURP* can be interpreted as the difference in PEAD between decile 10 and decile 1. A positive coefficient for *RSURP* suggests that PEAD increases with *SURP*. In the first set of regression, we classify firms by whether it is owned by any ETF or not. In the next set of regressions, we replace variable *ETF* by *SECTOR* and *ONLYBROAD*. *ETF*, *ONLYBROAD* and *SECTOR* are indicator variable defined identically to our firm-level analysis in section 4. Our variables of interest are the interaction terms, *RSURP*ETF*, *RSURP*SECTOR*, and *RSURP*ONLYBROAD.* A positive coefficient indicates a worsening of the drift while a negative coefficient indicates a mitigation in the drift.

We run the above regression model using two procedures – a pooled panel regression with two-way clustered t-statistics (clustered by firm and year), and quarterly regressions summarized using the Fama and MacBeth (1973) procedure. The results are presented in Table 9. The first column presents the pooled panel regression using *ETF*. Consistent with the presence of PEAD, we find that *RSURP* has a strong positive association with *POST60* with a coefficient of 0.094 (t-stat 9.51), suggesting the difference in PEAD between the firms with 10th decile of earnings

[16] The control variables are as follows. *SIZE* is the log of market value at the end of the fiscal quarter. *MTB* is the market-to-book ratio measured at the end of fiscal quarter. *BETA* is the estimated coefficient for market returns in the market model regression of a firm's daily returns on value-weighted market returns from all the trading days in the prior quarter. *PRERET* is the return momentum measured as the cumulated size-adjusted returns over the 20 trading days [-21,-2] before earnings announcements.

surprises and firms in the 1st decile of earnings surprises is about 9.4 percent. Consistent with prior research, we also find that the drift is negatively associated with size, positively associated with systematic risk and momentum and negative associated with size and market-to-book ratio. The interaction term *RSURP*ETF* has an insignificant negative coefficient (-0.007, t-stat -0.73). This suggests that overall ETF ownership does not mitigate the drift. The next column repeats the analysis using quarterly Fama and MacBeth (1973) regressions and finds similar results.

The next two columns of Table 9 repeat the analysis using the interaction of *RSURP* with *SECTOR* and *ONLYBROAD*. We find that sector ETF ownership is strongly associated with lower drift, while this effect is insignificant when ETF ownership is entirely non-sectoral. For the pooled specification, the coefficient on *RSURP*SECTOR* is -0.018 (t-stat -2.15), while the coefficient on *RSURP*ONLYBROAD* is 0.003 (t-stat 0.38). For the Fama and MacBeth specification, the coefficient on *RSURP*SECTOR* is -0.020 (t-stat -2.30), while coefficient on *RSURP*ONLYBROAD* is 0.001 (t-stat 0.16). This suggests that the mitigating effect of ETF ownership on PEAD seems to stem entirely from sector ETF ownership. These results are also consistent with our earlier results regarding the intra industry information transfer in our firm level analyses using Thomas and Zhang (2008) research design. Our findings suggest that sector ETF ownership/trading leads to the impounding of relevant information in the pre-earnings announcement period, which reduces the extent of PEAD.

7. Conclusion

This paper examines the role of ETFs in facilitating the flow of information between firms. Using earnings announcements as our information event, we examine the effect of ETF ownership on the flow of information between firms. We carry out this analysis in two distinct ways – at the firm level focusing on how ETF ownership affects intra-industry information transfers, as well as the ETF level focusing on how firms within the same ETF react to each other's earnings releases.

At the firm level, we find that the presence of ETF ownership of firms reduces the incidence of anomalous over-extrapolation of intra-industry information, especially for the case of sector ETF ownership. At the ETF level, we focus on the largest constituents of ETFs and separating them in the lead announcer and the follower firms. We find that while followers experience significant reaction to the earnings news of the leader, this is followed by a reversal, consistent with followers overreacting to the leader's earnings news. However, for sector ETFs, we find a much stronger initial reaction to the leader's earnings news followed by a much weaker reversal. Cross-sectional analyses suggest that the results are stronger when ETFs have high trading volume.

The emergence of ETFs has occurred in a time period that has also seen a number of significant changes influencing the capital markets including changes in market microstructure (decimalization), regulation (Sarbanes-Oxley), availability of peer information (EDGAR, XBRL) and functioning of market intermediaries such as analysts (global settlement). Hence ascribing causality to our results can be challenging. However, some of our additional analyses lend greater confidence to the explanation that the emergence of ETFs, especially sector ETFs, has influenced the transmission of information among firms. First, in our firm-level analysis, we find a mitigation of the anomalous over-extrapolation of industry information in the short window around the initiation of ETF ownership. Second, cross-sectional tests in our ETF level analysis suggest that the results are stronger when ETFs have high trading volume. Finally, when we parse out the correlation between the returns of leader and follower firms using path analysis, we find that the correlation is significantly mediated by ETF trading, especially for sector ETFs.

In our final tests, we examine whether the presence of ETF ownership mitigates the well-established anomaly associated with earnings announcements – the post-earnings announcement drift and find much lower drift in the presence of ETF ownership, but only for sector ETFs.

These results suggest that the answer to the question of whether ETFs help or hurt the flow of information between firms is contextual. ETFs have a bright side as well as a dark side. Markets seem to effectively use sector ETFs to transmit factor (industry) information impounded in earnings news but general broad market ETFs are not very useful (and potentially detrimental) to this type of information. Our results hence bridge the conflicting results documented in prior work on whether ETFs help or hinder market efficiency.

While our results seem to suggest that sector ETF have enhanced market efficiency, we need to caveat our finding that broad market ETFs may have hindered market efficiency. It is possible that broad market ETFs are not very effective in the context of earnings announcements, as these announcements often do not contain significant macro information. Market ETFs might be effective in other settings which we do not examine where macro information is being released and impounded by markets (e.g., GDP, interest rate or inflation data).

Our results broadly corroborate the findings of Huang, O'Hara and Zhong (2018) who find that industry ETFs help facilitate the hedging of industry specific risks by allowing traders to take long positions in individual stocks and short industry ETFs to offset industry exposure. They examine sector ETF short-selling behavior before earnings announcements, and find that pre-emptive trading prior to earnings announcements by industry participants (particularly hedge funds) reduces market reaction to earnings surprises. While our focus is on a different channel, in that we examine the role that ETFs (particularly sector ETFs) play in information transfer, the overall message of both papers is that industry or sector ETFs facilitate greater market efficiency.

References

Asher, H., 1983, Causal Modeling. Sage Publications Series: Quantitative Applications in the Social Sciences (QASS), 07-003. Beverly Hills, CA, and London, U.K.: Sage Publications.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2017, Do ETFs increase volatility? *Working paper, Ohio State University*.

Bernard, Victor L. and Jacob K. Thomas, 1989, Post-earnings announcement drift: delayed price response or risk premium, *Journal of Accounting Research* 27, 1–36.

Bernard, Victor L. and Jacob K. Thomas,1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305-340.

Bhattacharya, Ayan and Maureen O'Hara, 2017, Can ETFs increase market fragility? Effect of Information Linkages in ETF Markets, *Working paper, Cornell University*.

Bhattacharya, Nilabhra, Frank Ecker, Per M. Olsson, and Katherine Schipper, 2012, Direct and mediated associations among earnings quality, information asymmetry, and the cost of equity, *The Accounting Review* 87, 449-482.

Cong, Lin William and Douglas Xu, 2017, Rise of factor investing: asset prices, informational efficiency, and security design, *Working paper, University of Chicago*.

Da, Zhi and Sophie Shive, 2018, Exchange traded funds and asset return correlations, *European Financial Management* 24(1), 136-168.

Da, Zhi, Qianqiu Liu and Ernst Schaumburg. 2014. A Closer Look at the Short-Term Return Reversal. *Management Science* 60 (3), 658-674.

Fama, Eugene F. and James D. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607-636.

Fama, Eugene F. and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 42 (2), 153-193.

Freeman, Robert N. and Senyo Y. Tse, 1992, A nonlinear model of security price responses to unexpected earnings, *Journal of Accounting Research 30 (2), 185-209.*

Glosten, Lawrence, Suresh Nallareddy and Yuan Zou, 2017, ETF trading and informational efficiency of underlying securities, *Working paper, Columbia University*.

Huang, Mingyi, Xi Li and Shiheng Wang, 2015, Post-Earnings-Announcement Drift in global markets: evidence from an information shock, *The Review of Financial Studies* 28 (4), 1242-1283.

Huang, Shiyang, Maureen O'Hara and Zhuo Zhong, 2018, Innovation and informed trading: Evidence from industry ETFs, *Working paper, University of Hong Kong*.

ICI (Investment Company Institute) Fact Book, 2016, A review of trends and activities in the U.S. investment company industry, 56th edition, www.icifactbook.org.

Israeli, Doron, Charles M. C. Lee and Suhas A. Sridharan, 2017, Is there a dark side to exchange traded funds? An information perspective, *The Review of Accounting Studies* 22 (3), 1048-1083.

Leippold, Markus, Lujing Su and Alexandre Ziegler, 2016, How index futures and ETFs affect stock return correlations, *Working paper, University of Zurich.*

Beaver, William H., Maureen McNichols and Zach Z. Wang, 2017. Increased Information Content of Earnings Announcements in the 21st Century: An Empirical Investigation, *Working paper, Stanford University*.

Pan Kevin and Yao Zeng, 2017, ETF Arbitrage under liquidity mismatch, *Working paper, University of Washington.*

Subrahmanyam, Avanidhar, 1991, A theory of Trading in Stock index Futures, *The Review of Financial Studies* 4(1), 17-51.

Thomas Jacob and Frank Zhang, 2008, Overreaction to intra-industry information transfers? *Journal of Accounting Research* 46, 909-940.

APPENDIX

Variable Definitions

Panel A: Firm-Level Analysis

Variable	Definition
RESP	The average of firm's 3-day size-adjusted returns around its peers' earnings announcements, where the earnings announcement dates are at least five days prior to firm's earnings announcement date. The peer firm is defined as the firms from the same Fama French (1997) 48 industry.
ARET	Size-adjusted returns accumulated over 3 trading days starting from the firm's earnings announcement date.
ACC	Accruals measured as change in non-working capital less depreciation scaled by average total assets.
RET6	Buy and hold stock returns for the six month period up to one week before the firm's earnings announcement.
ERLYPRARET	Average of early peers' three-day announcement size-adjusted returns in the same quarter, where the peers' earnings announcements are at least five days prior to the firm's announcement.
ARET1	ARET lagged by one quarter
ARET4	ARET lagged by four quarters (same calendar quarter from prior year)
INST	Percentage of shares held by institutional investors
SIZE	Log of Market Capitalization at the end of prior fiscal year.
LOGBM	Log of the book-to-market ratio at the end of prior fiscal year.
ETFPERIOD	Indicator variable that equals 1 for periods 2002 and after and 0 otherwise
ETF	Indicator variable that equals 1 if cumulative holding by ETFs in a firm is greater than zero and 0 otherwise
ONLYBROAD	Indicator variable that equals 1 if all ETF ownership in a firm is non-sectoral and 0 otherwise
SECTOR	Indicator variable that equals 1 if cumulative holding by sector ETFs in a firm is greater than zero and 0 otherwise
SURP	Earnings Surprise defined as I/B/E/S actual earnings per share minus the last mean analyst consensus forecast before the earnings-announcement dates, scaled by the stock price at the start of the quarter.
RRDQ	Rank of Earnings release date for firms within a given Fama-French (1995) classification group for a given quarter minus 1. Equals 0 for the first firm, 1 for the second firm, 2 for the third firm, 3 for the fourth firm and 4 for the fifth firm.

Panel B: ETF-Level Analysis

Variable	Definition
LRET	Size-adjusted returns for the leader firm accumulated over 3 trading days starting from the day before leader's earnings announcement date
$FRET_{ANNC}$	Size-adjusted returns for the follower firm accumulated over 3 trading days starting from the day before leader's earnings announcement date
$FRET_{BETW}$	Size-adjusted returns for the follower firm accumulated starting 2 days after the leader's announcement date until 2 days before the follower's earnings announcement date
SEC	Indicator variable that equals 1 for sector ETFs and 0 for non-sector ETFs
HIGH	Indicator variable that equals 1 for leader-follower pair if the average trading volume over the 3 trading days starting from the day before leader's earnings announcement date exceed the median average daily ETF trading volume in the quarter before earnings season starts and 0 otherwise.

Panel C: Post-earnings announcement drift Analysis

Variable	Definition
POST60	Size-adjusted stock returns for the 60 day period after earnings.
ETF	Indicator variable that equals 1 if cumulative holding by ETFs in a firm is greater than zero and 0 otherwise
ONLYBROAD	Indicator variable that equals 1 if all ETF ownership in a firm is non-sectoral and 0 otherwise
SECTOR	Indicator variable that equals 1 if cumulative holding by sector ETFs in a firm is greater than zero and 0 otherwise
SURP	Earnings Surprise defined as I/B/E/S actual earnings per share minus the last mean analyst consensus forecast before the earnings-announcement dates, scaled by the stock price at the start of the quarter.
RSURP	Decile rank of SURP
SIZE	Log of market capitalization at the end of the fiscal quarter
BETA	Estimated coefficient for market returns in the market model regression of a firm's daily returns on value-weighted market returns from all the trading days in the prior quarter.
MTB	Market-to-book ratio measured at the end of fiscal quarter
PRERET	Return momentum measured as the cumulated size-adjusted returns over the 20 trading days [-21,-2] before earnings announcements.

Figure 1: Path Analysis

Figure 1a: Path Analysis for All ETFs

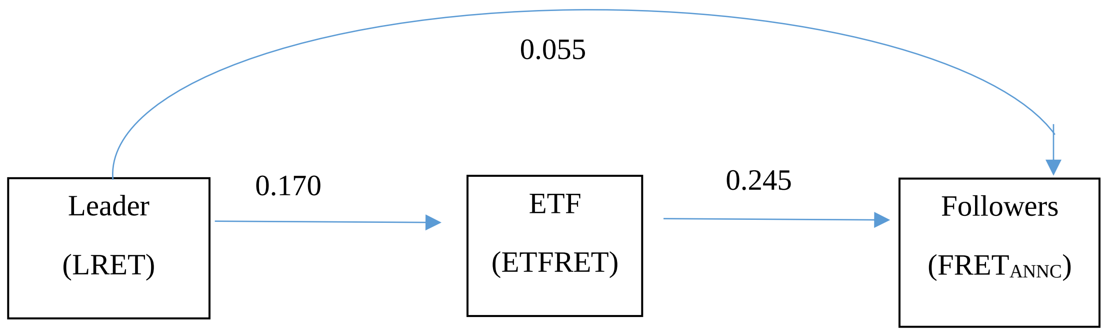

Figure 1b: Path Analysis for sector ETFs

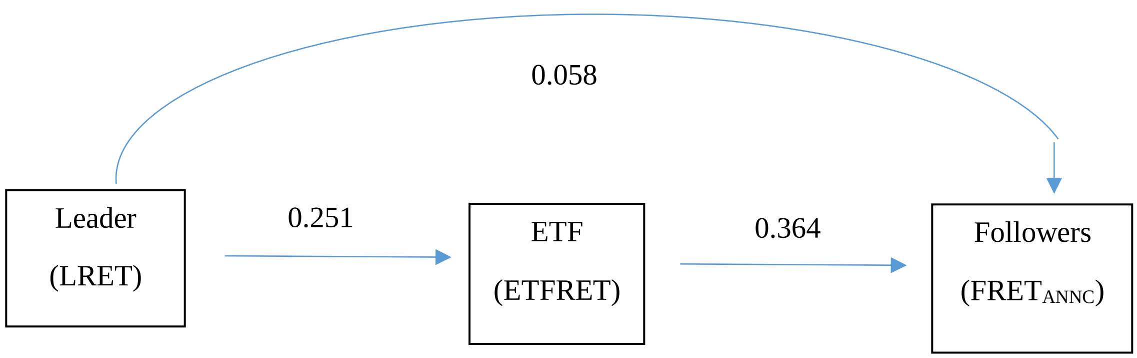

Figure 1c: Path Analysis for non-sector ETFs

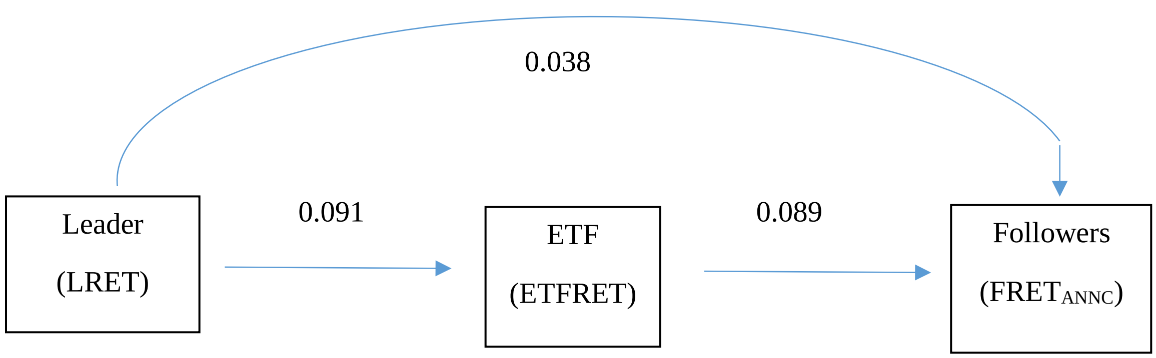

Figure 2: Trends in Earnings Response Coefficient across the Quarter

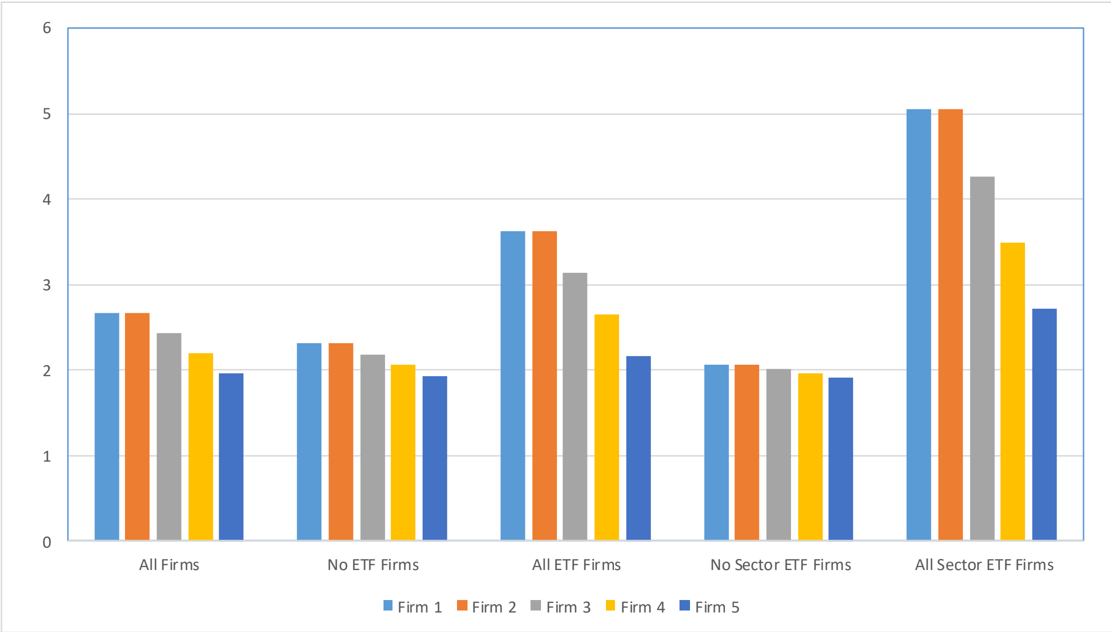

Figure 3: Trends in ETF ownership

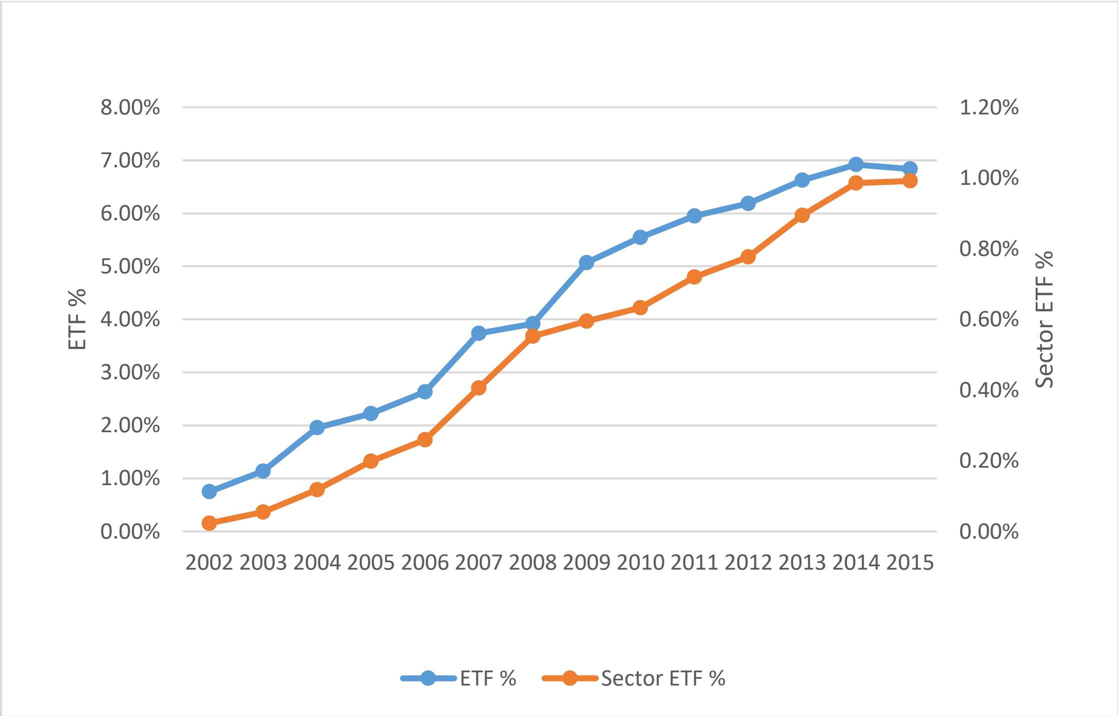

TABLE 1

Sample Selection and Distribution

Panel A: Sample Selection (ETF Level)

Sample Selection Criterion	*Observations*
Initial universe of ETF funds from CRSP as of 2015	2,091
Less: ETFs that are invested in stocks with no matches with Thomson-Reuters Mutual Fund Holding (S12) database	(1,604)
Number of distinct Equity ETFs with constituent holding information	487
Number of sector ETFs	214
Number of non-sector ETFs	273

Panel B: Distribution across Time

Year	*# distinct ETFs*
2002	106
2003	109
2004	138
2005	154
2006	177
2007	413
2008	473
2009	426
2010	453
2011	446
2012	434
2013	423
2014	407
2015	409

Panel C: Distribution of Sector ETFs by Sector

Sector	*Number of ETFs*	*% of Sector ETFs*
Aerospace	2	0.93%
Agriculture	1	0.47%
Banks	5	2.34%
Basic Materials	9	4.21%
Biotech	5	2.34%
Chemical	1	0.47%
Construction	3	1.40%
Consumer products	21	9.81%
Energy	10	4.67%
Environmental	2	0.93%
Financial Services	15	7.01%
Healthcare	29	13.55%
Industrials	9	4.21%
Infrastructure	2	0.93%
Internet	7	3.27%
Media	1	0.47%
Medical Devices	1	0.47%
Natural resources	2	0.93%
Nuclear	1	0.47%
Oil & Gas	8	3.74%
Pharmaceutical	4	1.87%
Precious Metals	2	0.93%
Real Estate	20	9.35%
Renewable Energy	4	1.87%
Retail	3	1.40%
Semiconductors	5	2.34%
Steel	1	0.47%
Technology	18	8.41%
Telecommunications	7	3.27%
Timber	1	0.47%
Transportation	1	0.47%
Utilities	10	4.67%
Water	4	1.87%
Total	214	100%

TABLE 2
ETFs and Investor Overestimation of Intra-Industry Information Transfers

This table considers a sample of 255,015 firm-quarters from 1985 to 2015. See the appendix for variable definitions. The dependent variable is ARET, which represents the 3-day size adjusted excess returns around earnings announcement. The independent variable of interest is RESP, the average of firm's 3-day size-adjusted returns around its peers' earnings announcements. Panel A runs the regressions for entire sample. Panel B runs the regression for a matched pair-sample using only the subset of data corresponding to firm that had ETF ownership sometime in the sample period. See section 4 for details of the research design. t-values are reported below each coefficient. All regressions are clustered at firm and year level. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, based on two-tailed tests.

Panel A: Intra-Industry Information Transfers in the pre-ETF and post-ETF periods

Model $ARET = \alpha_0 + \alpha_1{*}ETFPERIOD + \beta_1{*}RESP + \beta_2{*}RESP{*}ETFPERIOD + \beta_3{*}ACC + \beta_4{*}RET6 + \beta_5{*}ERLYPRARET + \beta_6{*}ARET1 + \beta_7{*}ARET4 + \beta_8{*}INST + \beta_9{*}SIZE + \beta_{10}{*}LOGBM + \varepsilon$

	Entire Sample	Pre-ETF Period	Post-ETF Period	Entire Sample
RESP	**-0.1950*****	**-0.2348*****	**-0.1488*****	**-0.2315*****
	(-10.78)	**(-10.01)**	**(-11.78)**	**(-10.00)**
RESP*ETFPERIOD				**0.0807*****
				(3.10)
ACC	0.0285***	0.0281***	0.0289***	0.0286***
	(10.06)	(7.59)	(6.26)	(10.02)
RET6	0.0066	0.018*	-0.0078	0.0066
	(0.81)	(1.89)	(-0.68)	(0.81)
ERLYPRARET	0.1078***	0.1077***	0.1067***	0.1073***
	(6.12)	(5.17)	(4.01)	(6.06)
ARET1	0.0226***	0.0313***	0.0152***	0.0227***
	(5.69)	(8.05)	(2.72)	(5.71)
ARET4	0.0015	0.0006	0.002	0.0015
	(0.60)	(0.13)	(0.62)	(0.58)
INST	0.0106***	0.0096***	0.0105***	0.0106***
	(10.75)	(5.29)	(8.12)	(10.83)
SIZE	-0.0007***	-0.0012***	-0.0002	-0.0007***
	(-3.60)	(-5.69)	(-0.63)	(-3.64)
LOGBM	0.0028***	0.0031***	0.0024***	0.0028***
	(5.26)	(3.43)	(4.30)	(5.24)
Adj. R^2	0.64%	0.87%	0.47%	0.65%
N	255,015	124,566	130,449	255,015

Panel B: Intra-Industry Information Transfers in the pre and post-ETF period for firms with ETF ownership

Model $ARET = \alpha_0 + \alpha_1{*}ETF + \alpha_1{*}ONLYBROAD + \alpha_2{*}SECTOR + \beta_1{*}RESP + \beta_2{*}RESP{*}ETF + \beta_3{*}RESP{*}ONLYBROAD + \beta_4{*}RESP{*}SECTOR + \beta_5{*}ACC + \beta_6{*}RET6 + \beta_7{*}ERLYPRARET + \beta_8{*}ARET1 + \beta_9{*}ARET4 + \beta_{10}{*}INST + \beta_{11}{*}SIZE + \beta_{12}{*}LOGBM + \varepsilon$

	Total ETF ownership	ETF ownership by ETF type
ETF	-0.0079***	
	(-3.85)	
ONLYBROAD		-0.0083***
		(-4.57)
SECTOR		-0.0072***
		(-2.72)
RESP	**-0.2434***** **(-9.13)**	**-0.2434***** **(-9.11)**
RESP*ETF	**0.0491** **(0.82)**	
RESP*ONLYBROAD		**0.0136** **(0.18)**
RESP*SECTOR		**0.1074***** **(3.07)**
ACC	0.0253***	0.0255***
	(2.88)	(2.87)
RET6	0.0150	0.0149
	(1.17)	(1.18)
ERLYPRARET	0.1267***	0.1248***
	(3.71)	(3.78)
ARET1	0.0319***	0.032***
	(6.66)	(6.65)
ARET4	0.0203***	0.0202***
	(2.64)	(2.63)
INST	0.0085***	0.0084***
	(6.04)	(5.78)
SIZE	-0.001	-0.0011
	(-1.23)	(-1.37)
LOGBM	0.0032	0.0032
	(1.52)	(1.54)
Adj. R^2	1.06%	1.07%
N	20,608	20,608

TABLE 3

Investors' Reaction to Earnings News for the Leader

The sample consists of leader-follower pairs within ETFs and covers years from 2002-2015. See Appendix for variable definitions. t-values are reported below each coefficient. All regressions are clustered at leader and leader's earnings announcement date level. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, based on two-tailed tests.

Panel A: Investors' reaction to follower firms on leader firm's earnings announcement

Model: $FRET_{ANNC} = \alpha_1 + \alpha_2*SEC + \beta_1*LRET + \beta_2*LRET*SEC + \varepsilon$

ETF period	All ETFs	Sector ETFs	Non-Sector ETFs	All ETFs
Intercept	-0.001***	0.000	-0.001***	-0.001***
	(3.25)	(0.48)	(4.34)	(4.34)
SEC				0.001**
				(2.40)
LRET	**0.049*****	**0.085*****	**0.023*****	**0.026*****
	(8.51)	**(7.95)**	**(3.69)**	**(3.69)**
LRET*SEC				**0.062*****
				(5.12)
N	31,992	16,281	15,711	31,992
Adj. R^2	0.93%	2.22%	0.25%	1.33%

Panel B: Investors' reaction partitioned by ETF trading volume

Model $FRET_{ANNC} = \alpha_1 + \alpha_2*SEC + \alpha_3*HIGH + \beta_1*LRET + \beta_2*LRET*SEC + \beta_3* SEC*HIGH + \beta_4*LRET*HIGH + + \beta_5*LRET*SEC*HIGH + \varepsilon$

	All pairs	High Volume pairs	Low Volume pairs	All pairs
Intercept	-0.001***	-0.002***	-0.001***	-0.001***
	(-4.34)	(-3.26)	(-3.05)	(-3.05)
SEC	0.001**	0.001*	0.001	0.001
	(2.40)	(1.82)	(1.64)	(1.64)
HIGH				0.000
				(0.52)
LRET	**0.023*****	**0.019****	**0.026*****	**0.026*****
	(3.69)	**(2.22)**	**(2.94)**	**(2.94)**
LRET*SEC	**0.062*****	**0.103*****	**0.017**	**0.017**
	(5.12)	**(5.46)**	**(1.21)**	**(1.21)**
SEC*HIGH				0.000
				(0.32)
LRET*HIGH				-0.007
				(-0.54)
LRET*SEC*HIGH				**0.086*****
				(3.68)
N	31,992	16,301	15,691	31,992
Adj. R^2	1.33%	2.53%	0.50%	1.59%

TABLE 4

Adjustment by Investors between Leader`s and Followers` Earnings Announcements

The sample consists of leader-follower pairs within ETFs and covers years from 2002-2015. See Appendix for variable definitions. t-values are reported below each coefficient. All regressions are clustered at leader and leader's earnings announcement date level. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, based on two-tailed tests.

Panel A: Adjustment by Investors between leader's and followers' earnings announcement

Model $FRET_{BETW} = \alpha + \beta_1 * FRET_{ANNC} + \beta_2 * SEC + \beta_3 * FRET_{ANNC} * SEC + \varepsilon$

	All ETFs	Sector ETFs	Non-Sector ETFs	All ETFs
Intercept	0.000	0.000	0.000	0.000
	(0.37)	(0.43)	(0.18)	(0.18)
$FRET_{ANNC}$	**-0.071*****	**-0.037*****	**-0.109*****	**-0.109*****
	(-3.89)	**(-1.89)**	**(-3.54)**	**(-3.54)**
SEC				0.000
				(0.18)
$FRET_{ANNC}$*SEC				**0.071****
				(1.99)
N	31,992	16,281	15,711	31,992
Adj. R^2	0.32%	0.08%	0.55%	1.34%

Panel B: Adjustment by Investors partitioned by ETF trading volume

Model: $FRET_{BETW} = \alpha_1 + \alpha_2 * SEC + \alpha_3 * HIGH + \beta_1 * FRET_{ANNC} + \beta_2 * FRET_{ANNC} * SEC + \beta_3 * SEC * HIGH + \beta_4 * FRET_{ANNC} * HIGH + \beta_5 * FRET_{ANNC} * SEC * HIGH + \varepsilon$

	All pairs	High Volume pairs	Low Volume pairs	All pairs
Intercept	0.000	0.000	0.000	0.000
	(0.18)	(0.52)	(0.22)	(0.22)
SEC	0.000	0.000	0.000	0.000
	(0.18)	(0.19)	(0.03)	(0.03)
HIGH				-0.000
				(-0.54)
$FRET_{ANNC}$	**-0.109*****	**-0.147*****	**-0.073****	**-0.073****
	(-3.54)	**(-2.86)**	**(-2.34)**	**(-2.34)**
$FRET_{ANNC}$*SEC	**0.071****	**0.104***	**0.045**	**0.045**
	(1.99)	**(1.83)**	**(0.99)**	**(0.99)**
SEC*HIGH				-0.000
				(-0.16)
$FRET_{ANNC}$*HIGH				-0.074
				(-1.25)
$FRET_{ANNC}$*SEC*HIGH				**0.059**
				(0.81)
N	31,992	16,301	15,691	31,992
Adj. R^2	0.32%	0.55%	0.15%	0.35%

TABLE 5
Path Analysis

This table reports the path analysis of how leader affects followers in underlying ETFs. This table consists 31,992 leader-follower pairs. See Appendix for variable definition. The results are illustrated in Figure 1. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, based on two-tailed tests.

	All ETFs	Sector ETFs	Non-Sector ETFs
Total effect			
ρ[LRET, $FRET_{ANNC}$]	0.097^{***}	0.149^{***}	0.050^{***}
	(17.43)	(19.42)	(6.25)
Direct path			
ρ [LRET, $FRET_{ANNC}$]	0.055^{***}	0.058^{***}	0.038^{***}
	(10.04)	(7.72)	(5.23)
Percentage	**57%**	**39%**	**84%**
Mediated path			
ρ [LRET, ETFRET]	0.170^{***}	0.251^{***}	0.091^{***}
	(31.31)	(11.43)	(11.48)
ρ [ETFRET, $FRET_{ANNC}$]	0.245^{***}	0.364^{***}	0.089^{***}
	(46.02)	(11.20)	(11.24)
Mediated effect	0.042^{***}	0.091^{***}	0.008^{***}
	(25.68)	(28.06)	(8.02)
Percentage	**43%**	**61%**	**16%**

TABLE 6
Falsification Analysis

This table repeats the analyses from Tables 3-4 using a sample of leader-follower firms in the pre-ETF time period from 1985 to 1999. The leader-followers are based on the pairs created in the ETF sample. See Appendix for variable definitions. t-values are reported below each coefficient. All regressions are clustered at ETF and year level. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, based on two-tailed tests.

Panel A: Investors' reaction to follower firms on leader firm's earnings announcement

Model: $FRET_{ANNC} = \alpha_1 + \alpha_2{*}SEC + \beta_1{*}LRET + \beta_2{*}FRET_{ANNC}{*}SEC + \varepsilon$

ETF period	All ETFs	Sector ETFs	Non-Sector ETFs	All ETFs
Intercept	0.0026***	0.0015***	0.0029***	0.0029***
	(10.76)	(4.37)	(10.97)	(10.97)
SEC				-0.001**
				(-3.35)
LRET	**0.081*****	**0.1314*****	**0.046*****	**0.046*****
	(12.80)	**(11.25)**	**(6.66)**	**(6.66)**
LRET*SEC				**0.086*****
				(6.32)
N	23,578	11,093	12,485	23,578
Adj. R^2	1.53%	3.51%	0.55%	1.98%

Panel A: Adjustment by Investors between leader's and followers' earnings announcement

Model $FRET_{BETW} = \alpha + \beta_1{*}FRET_{ANNC} + \beta_2{*}SEC + \beta_3{*}FRET_{ANNC}{*}SEC + \varepsilon$

	All ETFs	Sector ETFs	Non-Sector ETFs	All ETFs
Intercept	0.006***	0.005***	0.008***	0.008***
	(14.00)	(7.72)	(12.60)	(12.60)
SEC				-0.002**
				(-2.52)
$FRET_{ANNC}$	**-0.0284**	**-0.075****	**0.017**	**0.017**
	(-1.07)	**(-2.09)**	**(0.43)**	**(0.43)**
$FRET_{ANNC}$*SEC				**-0.092***
				(-1.73)
N	21,271	9,608	11,663	21,271
Adj. R^2	0.02%	0.15%	0.00%	0.11%

TABLE 7

Impact of ETFs on Earnings Response Coefficients across the Quarter

The sample consists of the five largest firms in each Fama and French (1997) industry grouping for which all data is available and comprises of 24,472 observations from 1985 to 2015. See the appendix for variable definitions. See section 6.1 for details of the research design. t-values are reported below each coefficient. All regressions are clustered at firm and year level. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, based on two-tailed tests.

Model $ARET = \alpha_0 + \beta_1*SURP + \beta_2*RRDQ + \beta_3*SURP*RRDQ + \beta_4*ACC + \beta_5*ARET1 + \beta_6*ARET4 + \beta_7*RET6 + \beta_8*INST + \beta_9*SIZE + \beta_{10}*LOGBM + \varepsilon$

	Entire Sample	No ETF Firms	All ETF Firms	No Sector ETF Firms	All Sector ETF Firms
SURP	**2.6634*** **	**2.3193*** **	**3.6338*** **	**2.0686*** **	**5.0511*** **
	(8.04)	**(5.25)**	**(11.39)**	**(5.00)**	**(8.25)**
RRDQ	-0.0004	-0.0008	0.0002	-0.0006	0.0002
	(-0.91)	(-1.04)	(0.40)	(-0.83)	(0.50)
SURP*RRDQ	**-0.2363*** **	**-0.1303**	**-0.491*** **	**-0.0537**	**-0.7772*** **
	(-2.32)	**(-1.29)**	**(-3.80)**	**(-0.55)**	**(-3.47)**
ACC	-0.0019	0.0076	0.0008	0.0062	0.0022
	(-0.22)	(0.70)	(0.06)	(0.57)	(0.12)
ARET1	0.0107	0.0444***	-0.0019	0.043***	0.0023
	(1.01)	(4.40)	(-0.11)	(4.42)	(0.16)
ARET4	0.0246***	0.0292	0.0272***	0.0319	0.0234***
	(2.66)	(1.18)	(3.27)	(1.36)	(2.48)
RET6	-0.0131	0.0212	-0.0368***	0.0156	-0.0400***
	(-0.69)	(0.58)	(-2.64)	(0.43)	(-2.79)
INST	-0.0073***	-0.0051	-0.0073***	-0.0052	0.0006
	(-3.09)	(-1.55)	(-3.04)	(-1.63)	(0.18)
SIZE	-0.0012***	-0.0012**	-0.0008*	-0.0012**	-0.0000
	(-3.88)	(-2.28)	(-1.70)	(-2.19)	(-0.01)
LOGBM	0.0000	0.0012	-0.0007	0.0012	-0.0008
	(-0.06)	(0.89)	(-0.88)	(0.85)	(-1.04)
Adj. R^2	2.28%	2.39%	3.07%	2.30%	3.46%
N	24,472	10,120	13,101	10,297	10,764

TABLE 8

Impact of ETF Ownership on Post-Earnings Announcement Drift: Portfolio Analysis

This table reports post earnings announcement drift (PEAD) by standardized unexpected earnings portfolios. Panel A outlines how the PEAD sample was generated. Panel B reports summary statistics for key firm characteristics. Panel C reports the returns in the 60 day period after earnings announcement (POST60) for deciles based on the earnings surprise (SURP). The first column presents the returns for the entire sample, the next two columns present the results for the subsample with ETF ownership below and above median respectively. Variables are defined in Appendix A. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, based on two-tailed tests.

Panel A; Drift Sample

Sample Selection Criterion	*Firm quarters*	*Distinct Firms*
All Compustat firms with fiscal quarter ending aligned with calendar quarter, incorporated in the US, and traded on NYSE NASDAQ and AMEX, and merged with Thomson-Reuters Mutual Fund Holding (S12) database to identify shares owned by ETF funds	204,915	6,809
Less: firm quarters with missing IBES earnings announcement date and timestamps and earnings forecasts from analysts	(65,803)	(1,416)
Firm quarters with earnings announcement dates and IBES information	139,112	5,393
Less: firm quarters with missing control variables	(5,141)	(84)
Final drift sample	133,971	5,309

Panel B: Summary Statistics (N=133,971)

Variable	*Mean*	*Median*	*P25*	*P75*	*Std*
ETF ownership	0.042	0.034	0.014	0.064	0.035
Sector ownership	0.005	0.001	0.000	0.004	0.013
Broad ETF ownership	0.037	0.031	0.013	0.056	0.029
Post60	0.003	-0.003	-0.085	0.080	0.165
Market Capitalization ($m)	5,269	845	268	2,812	20,777
Assets ($m)	10,939	1,126	300	4,011	78,333

Panel C: Post-Earnings-Drift by SURP Deciles

	All Firms	*ETF=1*	*SECTOR=1*	*ONLYBROAD=1*
SURP decile	*Post60*	*Post60*	*Post60*	*Post60*
1	-1.20%	-1.10%	-0.43%	-2.03%
2	-0.55%	-0.47%	-0.16%	-1.23%
3	-0.59%	-0.60%	-0.26%	-2.03%
4	-0.34%	-0.33%	-0.24%	-0.73%
5	-0.37%	-0.36%	-0.31%	-0.70%
6	0.19%	0.20%	0.18%	0.28%
7	0.41%	0.38%	0.15%	1.43%
8	0.99%	0.99%	0.70%	2.04%
9	1.42%	1.44%	1.19%	2.13%
10	3.14%	3.12%	2.79%	3.61%
(10)-(1)	4.34%***	4.23%***	3.22%***	5.64%***
	(16.47)	(15.62)	(9.82)	(12.60)
	Impact of ETF ownership	-0.11%	-1.12%***	1.30%
		(-0.29)	(-2.66)	(0.21)

TABLE 9

Impact of ETF Ownership on Post-Earnings Announcement Drift: Multivariate Analysis

This table reports regression with the dependent variable as the returns in the 60 day period after earnings announcement (POST60). Sample consists of 133,971 firm-quarters in the 2002-2015 period. See Appendix for variable definitions. Regressions are run either pooled with t-statistics controlling for two-way clustering by firm and year, or run quarterly and summarized using the Fama and MacBeth (1973) method. ***, **, * represent statistical significance at 0.01, 0.05 and 0.1 levels, respectively, using two-tailed tests. The first two columns present regressions using the below model. In the next two columns, we replace ETF with SECTOR and ONLYBROAD.

$$POST60 = \alpha + \beta_1{*}RSURP + \beta_2{*}ETF + \beta_3{*}RSURP{*}ETF + \beta_4{*}SIZE + \beta_5{*}BETA + \beta_6{*}MTB + \beta_7{*}PRERET\ \beta_8{*}RSURP{*}SIZE + \beta_9{*}RSURP{*}BETA + \beta_{10}{*}RSURP{*}MTB + \beta_{11}{*}RSURP{*}PRERET + \varepsilon.$$

	Pooled	*Fama-Macbeth*	*Pooled*	*Fama-Macbeth*
Intercept	-0.039***	-0.043***	-0.027***	-0.032***
	(-7.13)	(-7.02)	(-4.74)	(-4.54)
RSURP	**0.094*** **	**0.094*** **	**0.076*** **	**0.074*** **
	(9.51)	**(9.49)**	**(7.25)**	**(6.79)**
ETF	0.008*	0.009		
	(1.81)	(1.57)		
SECTOR			0.016***	0.017**
			(3.36)	(2.34)
ONLYBROAD			0.002	0.005
			(0.40)	(0.79)
RSURP*ETF	**-0.006**	**-0.007**		
	(-0.73)	**(-0.89)**		
RSURP*SECTOR			**-0.018** **	**-0.020** **
			(-2.15)	**(-2.30)**
RSURP*ONLYBROAD			**0.003**	**0.001**
			(0.38)	**(0.16)**
SIZE	0.003***	0.003***	0.001	0.001
	(4.87)	(3.33)	(0.80)	(0.59)
BETA	-0.006	0.008	-0.008	0.007
	(-0.89)	(0.34)	(-1.06)	(0.30)
MTB	-0.000	-0.000	0.000	-0.000
	(-0.10)	(-0.29)	(0.25)	(-0.21)
PRERET	-0.029**	-0.025	-0.029**	-0.025
	(-2.48)	(-1.01)	(-2.47)	(-1.01)
*RSURP*SIZE*	-0.008***	-0.008***	-0.004***	-0.004***
	(-8.08)	(-7.01)	(-3.69)	(-2.76)
*RSURP*BETA*	0.010	0.006	0.012	0.007
	(0.88)	(0.39)	(1.05)	(0.450)
*RSURP*MTB*	-0.001	-0.001	-0.001*	-0.001
	(-1.46)	(-1.00)	(-1.76)	(-1.03)
*RSURP*PRERET*	0.033*	0.018	0.032*	0.017
	(1.67)	(0.95)	(1.66)	(0.88)
Adj. R^2	0.57%	3.32%	0.60%	3.93%