Innovation and Informed Trading: Evidence from Industry ETFs

Shiyang Huang, Maureen O'Hara, and Zhuo Zhong*

January 2018

We hypothesize that industry exchange traded funds (ETFs) encourage informed trading on the underlying firms through facilitating the hedge of industry-specific risks. We find that industry ETF memberships increases hedge funds' abnormal holdings before earnings announcements and reduces the market reaction to the firm's earnings surprise, especially the positive surprise. In addition, we show that short interest on industry ETFs positively predicts returns on these ETFs and the percentage of positive earnings announcements of underlying firms. Our results suggest that financial innovations such as industry ETFs can be beneficial for informational efficiency if they help investors to hedge risks.

* Shiyang Huang (huangsy@hku.hk) is at University of Hong Kong; Maureen O'Hara is at the Johnson College of Business, Cornell University and UTS; Zhuo Zhong (zhuo.zhong@unimelb.edu.au) is at University of Melbourne. The online appendix is available at: https://www.dropbox.com/s/vyl9hgah9ufal57/online_appendix_Jan27.pdf?dl=0.

Innovation and Informed Trading: Evidence from Industry ETFs

January 2018

We hypothesize that industry exchange traded funds (ETFs) encourage informed trading on the underlying firms through facilitating the hedge of industry-specific risks. We find that industry ETF memberships increases hedge funds' abnormal holdings before earnings announcements and reduces the market reaction to the firm's earnings surprise, especially the positive surprise. In addition, we show that short interest on industry ETFs positively predicts returns on these ETFs and the percentage of positive earnings announcements of underlying firms. Our results suggest that financial innovations such as industry ETFs can be beneficial for informational efficiency if they help investors to hedge risks.

Innovation and Informed Trading: Evidence from Industry ETFs

Few financial innovations in recent times have had the impact of exchange traded funds (ETFs). With assets approaching $3.5 trillion, ETFs are now larger than hedge funds. Worldwide there are approximately 5000 exchange traded funds, making ETFs the preferred investment approach for a wide range of investors. Indeed, it does not seem an exaggeration to argue that the growth of passive investing via ETFs has posed a "disruptive innovation" for the entire asset management industry.[2] For many investors, the main innovation of ETFs is to provide a more liquid, lower-cost alternative to mutual funds. For others, the innovation is access to previously unavailable asset classes. In this paper, we argue that another, perhaps under-appreciated, innovation is an expanded ability to hedge. We demonstrate that this aspect of ETF innovation has a direct impact on the nature of informed trading and the efficiency of the market.

We hypothesize that ETFs can reduce hedging costs for informed investors. We develop this hypothesis based both on the theoretical literature on financial innovation and on industry reports on ETFs. The literature shows that financial innovations such as introducing a new security can improve risk-sharing (Allen and Gale (1994)). Moreover, Dow (1998) and Simsek (2013) show that the new security could enhance investors' arbitrage profits if it could be used to hedge their arbitrage risk. This idea of hedging is also widely observed in practice, especially in reports on how hedge funds use ETFs. For example, Bloomberg recently reported "Hedge funds mainly use ETFs to take short positions. … As a group, hedge funds have $105 billion in short ETF positions — more than double their $43 billion in long positions. … The funds'

[2] Ananth Madhavan makes the case for such disruptive innovation in his book Exchange Traded Funds and the New Dynamics of Investments, (FMA Survey and Synthesis: 2016).

shorts don't necessarily indicate bearish sentiment, but rather are used to hedge out part of the market in order to isolate a long position." [3]

To investigate this hedging role, we focus on the role played by industry ETFs. Industry ETFs are appealing for two reasons. First, it is natural to use an industry ETF to hedge the industry risk of a firm. Because a firm's risk usually includes market risk, industry risk, and firm-specific risk, an informed trader hoping to profit from firm-specific information will want to hedge the market risk and the industry risk. While index futures or index ETFs are used to hedge the market risk, the advent of industry ETFs provides a vehicle to hedge the industry risk. Second, from a technical perspective, the cross-sectional variation of the industry ETF membership and the time-series variation of the inception date of each industry ETF allow us to identify and quantify the economic impact of the industry ETF. To our knowledge, ours is the first paper to address these industry-hedging effects specifically.

We first establish two important facts. We show that the industry ETF is more likely to experience large short interest than either the non-industry ETF or individual stocks. Indeed, we find that the industry ETF has a short interest ratio (short interest/shares outstanding) of 60% at the 95 percentile, in contrast to the non-industry ETF (individual stocks) which has less than 20% (17%) at the same percentile.[4] More interesting, we find that large short interest on the industry ETF does not imply a bearish outlook. On the contrary, we find that large short interest predicts more *positive* earnings surprises among the underlying stocks of the industry ETF.

These findings are consistent with the hypothesis that informed investors use the industry ETF to hedge their long positions on firms with positive private information, implementing in effect a "*long the underlying, short the ETF*" strategy. Our hedging hypothesis generates two

---

[3] See Bloomburg Intelligence, September 8th 2017.

[4] Stock level short interest is reported in the online appendix (Table A.6).

implications. First, the ability to hedge with the industry ETF incentivizes informed investors to trade more aggressively, as information rents become easier to capture. Consequently, the market becomes more informationally efficient. Second, short interest on the industry ETF generates a temporary price impact leading to a price reversal. The price reversal means short interest could positively predict the return of the industry ETF.

We test the first implication with earnings announcement events. We conjecture that earnings announcements are associated with lower market reactions if informed investors trade beforehand and prices incorporate more information before the announcement. Consistent with this conjecture, we find that industry ETF membership significantly reduces the market reaction to positive earnings surprises. We further demonstrate that industry ETF membership leads to *ex-ante* more aggressive buying by hedge funds when the future earnings surprise is positive. On the other hand, industry ETF membership does not generate any impact if the future earnings surprise is negative.

We test the second implication regarding the return predictability of short interest on the industry ETF by using a Fama-MacBeth approach. We find that the change in the short interest ratio (ΔSIR) positively predicts the future return for an industry ETF, whereas we find the opposite pattern for short interest on the firms in the industry ETF. That is, at the member-firm level, ΔSIR negatively predicts the future return. The firm-level result is consistent with past studies (see Rapach, Ringgenberg, and Zhou (2016)), but as we demonstrate short interest on the industry ET itself behaves differently from short interest on member firms. Our result is consistent with the hedging hypothesis that extreme short interest reflects the long/short strategy from informed investors rather than bearish speculation on the industry.

Our paper contributes to the literature in a number of areas. The financial innovation literature shows that an important motive for creating a new financial security is to allow investors to hedge substantial risks, or more generally, to complete the financial market. Duffie and Rahi (1995) provide a comprehensive survey on this topic. Completing the financial market enables investors to better span their investment opportunities. Investors can isolate some risks with the new financial security which could lead to more beneficial portfolios (Chen (1995)). It could also lead to more informed trading. Dow (1998) shows that informed investors with better hedging trade more aggressively on information. Simsek (2013a, 2013b) argues that financial innovation leads to more speculation among investors with different beliefs on different risk factors, as hedging becomes easier. Recently, a small theoretical literature studies the impact of the ETF per se. Cong and Xu (2016) show that introducing composite securities facilitates trading common factors in assets' liquidation values. Bhattacharya and O'Hara (2017) study how inter-market information linkages in ETFs can exacerbate market instability and herding. We provide direct empirical evidence showing that financial innovation, in particular the industry ETF, is associated with more aggressive informed trading from sophisticated investors such as hedge funds.

Our paper also contributes to the growing empirical ETF literature. Past studies center around the price impacts of arbitrage activity between the ETF and its underlying. Ben-David, Franzoni, and Moussawi (2014) find the ETF arbitrage activity increases non-fundamental volatility on underlying stocks. Wermers and Xue (2015) and Madhavan and Sobczyk (2015) also find ETFs are associated with higher volatility of underlying. While the literature seems to agree on that the ETF increases volatility in the underlying, the impact on liquidity and informational efficiency remains undetermined. Madhavan and Sobczyk (2015) find that ETFs have heterogeneous effects on price efficiency of underlying assets and the effect depends on the liquidity of ETFs. Glosten, Nallareddy, and Zou (2017) find ETF's membership positively affects informational efficiency at the stock level, especially, the incorporation of earnings information. On the other hand, Israeli, Lee, and Sridharan (2017) find the ETF ownership is associated with a larger bid-ask spread and less informative price. Our paper provides a new perspective to study the ETFs' impacts on informational efficiency. We focus on industry ETFs and show that, by facilitating informed traders' hedging needs, industry ETFs encourage more informed trading.

Last but not least, our paper adds to the literature on short selling. Although past empirical findings largely imply short sellers have superior information in predicting future abnormal returns (see Desai et al. (2002), Asquith, Pathak, and Ritter (2005), Boehmer, Jones, and Zhang (2008), and Diether, Lee, and Werner (2009)), the literature also acknowledges some short sellers are merely hedgers. Hedgers short a stock to hedge their positions in other assets, for example, the delta hedge in put option trading. Ofek, Richardson, and Whitelaw (2004) show short-sale constraints in the underlying stock increase violations of put-call parity, suggesting that difficulty hedging an option position affect prices in both the stock and option market. Battalio and Schultz (2011) and Grundy, Lim, and Verwijmeren (2012) study the 2008 short-sale ban and find that option bid-ask spreads increase for banned stocks. Their results imply option markets are disrupted when hedging the underlying stock with short-sales becomes difficult (or almost impossible). Our analysis of short interest on industry ETFs provides additional evidence on the impact of hedging-based short-sales. Our results also show an intriguing asymmetry as the hedging effect we identify only operates with respect to positive news, and not with negative news. We conjecture that this is due to the higher short-sale costs of individual stocks, making a "*short the stock, long the ETF*" strategy uneconomic.

This paper is organized as follows. The next section sets out the data and sample statistics. Section 2 then investigates the role of short interest in industry ETFs, examines its relation to underlying firms' earning surprise, and estimates its impact on the market reaction to earning surprises. We also provide evidence on the channel through which these effects operate by showing how industry ETFs change hedge funds trading behavior. In Section 3 we test the relation between the short interest ratio and returns in industry ETFs using a Fama-MacBeth approach. Section 4 provides some additional robustness and placebo tests. Section 5 is a conclusion.

1. Data description and sample statistics

Our sample consists of two sets of data. The first data set is the industry ETF sample. For each U.S. industry ETF, we track its short interest, holdings, price, and volume from its inception until December 2016. The second data set is the earnings announcement sample. This data contains the earnings announcements of all publicly listed firms from January 1995 to December 2016. We also collect a variety of related data such as hedge fund holdings, mutual fund holdings, and firm characteristics. In this section, we discuss the construction of these two data sets in detail.

1.1 The ETF level data

A. The equity ETF

To construct the list of industry ETFs on US equity, we first need a list of US equity ETFs. We start with the fund universe of the CRSP Survivor-biased-free Mutual Fund database ranging from July 2003 to December 2016. We identify a fund as an ETF if the "et_flag" of the fund is "F." Also, we require these funds to have the CRSP share-code of either "44" or "73." To obtain the non-synthetic US equity ETF, we drop funds whose name contains "bond," "bear," or "hedged." And to ensure our list consists of only equity ETFs, we apply a filter which requires our sample ETFs to have at least 80% investment in US common domestic stocks. After those steps, we merge our list with a snapshot of all US equity ETFs from ETFDB in July 2017.[5] Our final sample consists of 449 US equity ETFs, which is close to past studies.[6] For each ETF, we track its holdings information since the inception date.[7]

B. The industry ETF

We extract industry ETFs from the abovementioned equity ETFs based on the holdings information. We match an ETF's holdings with the Fama-French 12 industry classification, and then identify the industry in which the ETF has the most investment. To qualify for an industry ETF, we require the dominating industry investment exceeds one-third of the ETF's portfolio size. This gives us 217 industry ETFs. We filter out ETFs whose name contains "value," "growth," "Russell," "dividend," or "momentum" to ensure the ETF targets for the industry coverage. After this step, we are left with 150 ETFs. We further require that the ETF consists of at least 30 stocks (in the on-line Appendix we remove this requirement and show that our results hold with this less restrictive industry ETF list). Finally, we obtain a list of 116 industry ETFs covering 11 out of 12 industries in the Fama-French classification. Figure 1 shows the time series growth of the total net asset value and industry coverage of our industry ETF sample.

[Insert Figure 1 Here]

C. The price, volume, and short interest data for equity ETFs

We obtain the monthly price and volume for our ETF sample from CRSP. The monthly short interest for both equity ETFs and their underlying firms are from COMPUSTAT. Panel A and B in Table 1 report the summary statistics of price, volume, and short interest for our ETF sample. We report the industry and non-industry ETF, respectively.

[Insert Table 1 Here]

[5] ETFDB is a website providing detail information on ETFs, see www.etfdb.com for details.

[6] Glosten, Nallareddy, and Zou (2017) identify 447 ETFs between 2004 and 2013; Israeli, Lee, and Sridharan (2017) identify 443 ETFs between 2000 and 2014; Da and Shive (2015) identify 549 ETFs between 2006 and 2013; Li and Zhu (2017) identify 343 ETFs from 2002 to 2013.

[7] We use 13F data from Thompson Reuters for fund holdings, and complement it with the CRSP Survivor-biased-free Mutual Fund database.

1.2. The firm level data

A. Data on earnings announcements

We construct our data on earnings announcements based on analyst-target-price forecasts from the Institutional Brokers' Estimate System (I/B/E/S), quarterly financial statements from COMPUSTAT, and financial market data from CRSP. We focus on quarterly earnings announcements that are available in both COMPUSTAT and I/B/E/S.[8] Following Livnat and Mendenhall (2006) and other papers in this literature, we impose the following restrictions:

(1) Ordinary common shares listed on the NYSE, AMEX, or NASDAQ.

(2) The earnings announcement date is reported in both COMPUSTAT and I/B/E/S, and the earnings report dates in COMPUSTAT and in I/B/E/S differ by no more than one calendar day.

(3) The price-per-share at the end of the fiscal quarter is available from COMPUSTAT and is greater than \$1.

(4) The market value of equity at the fiscal quarter-end is available and is larger than \$5 million.

(5) Daily stock returns are available in CRSP for the dates around the earnings announcement. Moreover, we should be able to assign stock to one of the six Fama-French benchmark portfolio based on size and book-to-market ratio.

In the analysis of the market reaction to the earnings announcement, we define an earnings surprise by the standardized unexpected earnings (SUE). The SUE of firm i at quarter t is calculated as $SUE_{i,t} = \frac{EPS_{i,t}-EPS_{i,t-4}}{\sigma_{i,t}}$, where EPS_t is the earnings per share at quarter t, and

[8] We use the link table provided by Prof. Byoung-Hyoun Hwang from Cornell University. This link table provides a mapping from I/B/E/S ticker to CRSP permno and can be downloaded from his personal webpage: http://www.bhwang.com/code.html.

EPS_{t-4} is the earnings per share at the same quarter in the previous year. $\sigma_{i,t}$ is the standard deviation of $EPS_{i,t} - EPS_{i,t-4}$ in the last eight consecutive quarters.

B. The hedge fund and mutual fund list

We construct a list of hedge funds based on Form ADV (an SEC regulatory filing). After 2011, all U.S. hedge fund advisers with more than $150 million in asset under management are required to file Form ADV. Following Brunnermeier and Nagel (2005), Griffin and Xu (2009), and Jiang (2017), we take two steps to construct the hedge fund list. First, an asset management adviser from Form ADV is identified as a hedge fund if 80% of its assets are in the hedge fund business (as reported in Form ADV). Second, the list of hedge funds in the first step is manually merged with Form 13 (CDA/Spectrum) through advisers' names.[9] The CDA/Spectrum database is also used to construct the list of US equity mutual funds. Following Lou (2012), mutual funds in our sample have a minimum fund size of $1 million, and the total net assets, TNA, reported by the CDA/Spectrum do not differ from CRSP's TNA by more than a factor of 2. The equity mutual funds in our sample have investment objective codes: aggressive growth, growth, growth and income, balanced, unclassified, or missing.

C. Data on institutional holdings

We follow Ben-David, Franzoni, and Moussawi (2011) to construct institutional holdings for each firm at each quarter based on the Thompson Reuters 13F data.[10] Merging this with the above list on hedge funds, we obtain hedge fund holdings on our sample firms at each quarter. To estimate abnormal holdings, we take the difference between the current quarter holdings and the moving average of the past four quarters holdings. Similarly, we obtain mutual fund abnormal holdings on our sample firms at each quarter. In Panel C Table 1, we report summary

[9] The detailed description can be referred to the online appendix of Jiang (2017).
[10] WRDS provides the detail code in constructing institutional holdings from the 13F data, see https://wrds-web.wharton.upenn.edu/wrds/research/applications/ownership/Institutional%20Trades/.

statistics for our earnings announcements sample after winsorizing the bottom and top 1%. All our variables have a distribution similar to past studies.[11]

2. Industry ETFs, information, and hedging

Can industry ETFs facilitate informed trading and enhance the efficiency of the market? In this section, we address this question by first examining the behavior of short interest on industry ETFs with a focus on whether it reflects speculation or hedging. We then look at how the level of short interest affects the market reaction to the earnings announcement. We further investigate the channels by which such effects arise by looking at the impact of industry ETFs on the portfolio holdings of hedge funds and mutual funds.

2.1 Why do investors short industry ETFs?

We begin our analysis by investigating the properties of short interest in ETFs. Panels A and B in Table 1 show that the short interest ratio of industry ETFs has a much longer right tail than that of non-industry ETFs. The industry ETF has a short interest ratio of 60% at the 95th percentile whereas the latter has less than 20% at the same percentile. Figure 2 shows the histogram of the short interest ratio. For industry ETFs, we observe a significant concentration of the short interest ratio at 100%. Such a pattern is not observed among non-industry ETFs. The longer right tail of the short interest ratio indicates that industry ETFs experience more extreme short positions.[12]

[Insert Figure 2 here]

One natural explanation for the pattern in the industry ETF is that investors are betting against a specific industry, e.g., investors shorting the financial industry during the 2008

[11] Our hedge fund abnormal holdings has similar magnitude on the mean and standard deviation as Chen, Da, and Huang (2016).

[12] In constructing the short interest ratio, we replace all ratios above 100% with 100%. In other words, the concentration of the short interest ratio at 100% represents a large cumulative mass of short interest exceeding 100%.

financial crisis. We call this the speculation hypothesis. An alternative hypothesis captures the hedging motive.[13] Informed investors short an industry ETF to hedge their long positions on particular underlying firms for which they have private information on firm-specific fundamentals. This "*long-the-underlying*" and "*short-the-industry-ETF*" strategy enables informed investors to hedge their industry risk exposure while obtaining rewards for certain individual stocks (in that industry).

These two hypotheses have distinct predictions regarding the future outlook of an industry ETF. The speculation hypothesis predicts a bearish outlook of the industry ETF, whereas the hedging hypothesis offers the opposite prediction: The large short position on an industry ETF reflects many informed investors with optimistic information about an underlying firm hedging heavily to isolate their positions from the industry risk.

To test these two hypotheses, we construct a quarterly measure that captures the earnings performance of each ETF's underlying. First, we compute the ratio of positive earnings among underlying firms in an ETF. We define a firm to have positive earnings if its SUE is in the top 25% of the entire sample. Second, at every quarter, we compute the positive earnings ratio as the ratio of underlying firms in an ETF that have positive earnings. This positive earnings ratio measures the economic outlook of an ETF's underlying. Panel D of Table 1 reports the summary statistics of our positive earnings ratios for industry and non-industry ETFs.

After the above construction, we use the following regression to test predictions based on the speculation and hedging hypotheses:

$$PosSUE_{i,t} = \alpha_i + \alpha_t + \beta_1 SIR_{i,t-1} + controls + \epsilon_{i,t}, \tag{1}$$

[13] Note that we do not view these hypotheses as mutually exclusive. Some traders may use industry ETFs to speculate, others to hedge. Our interest here is to determine any empirical linkages of these industry ETFs to informed trading.

$$PosSUE_{i,t} = \alpha_i + \alpha_t + \beta_1 SIR_{i,t-1} + \beta_2 DummyIndetf \times SIR_{i,t-1} + controls + \epsilon_{i,t}, \tag{2}$$

where $PosSUE_{i,t}$ is the positive earnings ratio for ETF i at quarter t, and $SIR_{i,t-1}$ is the lagged short interest ratio for i. We include the log total net asset value of i in the contemporaneous quarter as a control variable. We also control for year, quarter, and ETF fixed effects. Standard errors are clustered by ETFs and quarters. We estimate equation (1) on industry and non-industry ETFs, respectively, and estimate equation (2) on all ETFs. In equation (2), the dummy variable $DummyIndetf$ interacting with the lagged short interest ratio captures the different predicting power of the short interest ratio between industry and non-industry ETFs. We show the regression result in Table 2.

[Insert Table 2 here]

For an industry ETF, we find that a large short interest ratio predicts more positive earnings among its underlying firms. We find the opposite (or no predicting power) for a non-industry ETF. Our regression results from equation (2) suggest that the predicting power of the short interest ratio is significantly different between industry and non-industry ETFs. This difference becomes even more pronounced if we exclude 2007 and 2008 crisis periods from our sample (see Panel B of Table 2).[14]

These regression results are consistent with the hedging hypothesis. Large short interest predicting more positive earnings is consistent with the long/short strategy carried out by informed investors, who are long the underlying based on their positive private information. More intriguing, we only observe this predictability among industry ETFs, and not among non-industry ETFs.

[14] We have also tested the predictability of the industry ETF's short interest on the negative earnings ratio. The negative earnings ratio is constructed similar to the positive earnings ratio with the negative earnings defined as the SUE in the bottom 25% of the sample. We do not find the industry ETF's short interest has significant predicting power on the negative earnings ratio.

We do not rule out the speculation hypothesis. In fact, we find that when including the financial crisis into our sample period, the predictability of past short interest is reduced by over 30% (from 0.0436 to 0.0324). The reduction is possibly due to high short interest and poor performance of the financial sector during the crisis. Hence, to focus on the hedging hypothesis and its economic implications, we exclude 2007 and 2008 in the remaining analysis. Results from the entire sample, showing that our results hold more generally over the whole sample period, are available in an online appendix.

2.2 The market reaction to the earnings surprise

In the previous sub-section, we find evidence consistent with informed investors using industry ETFs to hedge their long positions on the corresponding underlying. Such a strategy seems feasible as Li and Zhu (2017) show that ETFs are relatively easy to short, and the growing popularity of industry ETFs suggests it is becoming less expensive to do so. Reducing the costs of hedging facilitates informed investors trading which, in turn, should make the market more informative. Therefore, we hypothesize that the industry ETF has a positive impact on the informational efficiency of the market.

To test this hypothesis, we focus on earnings announcement events. We examine the market reaction to the earnings surprise, and study if industry ETF membership affects the reaction. More specifically, we run the regression:

$$CAR_{i,t} = \alpha_i + \alpha_t + \theta_1 SUE_{i,t} + \theta_2 DummyIndetfown + \theta_3 DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}. \quad (3)$$

$CAR_{i,t}$ is the -1 to +1 cumulative abnormal daily return around the earnings announcement date based on the Fama-French three factor model.[15] $SUE_{i,t}$ is the standardized earnings surprise. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an

[15] Our result remains the same when we use the Fama-French four factor (including the momentum factor) model.

industry ETF. The interaction term, $DummyIndetfown{\times}SUE_{i,t}$, captures how industry ETF membership affects the relationship between the market reaction ($CAR_{i,t}$) to the earnings surprise ($SUE_{i,t}$). In our controls, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts (see Panel C Table 1 for summary statistics on the variables used in Eq.(3)). In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates.

[Insert Table 3 here]

We estimate equation (3) on the full sample of our earnings announcements ranging from 1995 to 2016, excluding 2007 and 2008.[16] We also divide the earnings announcements sample into "Negative SUE" and "Positive SUE" groups. The "Negative SUE" group consists of the bottom 25% SUE of our sample, while the "Positive SUE" group consists of the top 25% SUE of our sample. We report the regression results in Table 3.

Our regression results support the hypothesis that the industry ETF enhances pricing efficiency of its underlying. We find that the market reacts less to an earnings surprise when a firm is the constituent of an industry ETF. The coefficient on $DummyIndetfown{\times}SUE_{i,t}$is significantly negative, implying that more information is incorporated into the market before the earnings announcement. This could be due to more informed trading on firms in industry ETFs before earnings announcements. Additionally, the coefficient is only significantly negative when there is a positive earnings surprise. As the positive earnings surprise indicates the positive firm-specific information *ex-ante*, this finding further substantiates our hypothesis. We hypothesize that the industry ETF encourages more informed trading through facilitating the "*long-the-underlying*" and "*short-the-industry-ETF*" strategy. As this strategy is applicable

[16] All our results are similar when we include 2007 and 2008. Please see the online appendix for more details.

when there is positive information on the underlying firm, the impact of the industry ETF emerges under positive information rather than negative information.

2.3. Hedge funds abnormal holdings and the earnings surprise

We showed that the market reacts less to an earnings surprise when the firm is a member of an industry ETF. With more informed investors trading in advance, this suggests that the market becomes more informationally efficient. There could be, however, other hypotheses explaining the decreasing reaction; for example, one could argue that it reflects the market's slow adjustment to the earnings news. To investigate more thoroughly if the industry ETF encourages more informed trading, and hence, improves efficiency, we consider the potential channels of our hedging argument. As noted earlier, hedge funds are active users of short interest ETF strategies and so seem likely candidates to implement this hedging strategy. To explore this possibility, we study hedge fund portfolio holdings.

Using data on aggregate hedge fund holdings, we run the regression model:

$$\begin{aligned} HfAhdngRatio_{i,t}, & \\ &= \alpha_i + \alpha_t + \theta_1^H SUE_{i,t} + \theta_2^H DummyIndetfown \\ &+ \theta_3^H DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}. \end{aligned} \qquad (4)$$

$HfAhdngRatio_{i,t}$ is the preceding abnormal holdings by hedge funds on firm i at time t. The abnormal holdings variable is estimated as the difference between the current quarter holdings and the moving average of holdings in past four quarters, standardized by the total shares outstanding. The summary statistics of $HfAhdngRatio_{i,t}$ are reported in Panel C Table 1. Other variables are the same as in equation (3).

In equation (4), we treat the current quarter earnings surprise, $SUE_{i,t}$, as the realization of firm-specific information. If hedge funds' abnormal holdings (preceding the quarter) load significantly on $SUE_{i,t}$, it implies hedge funds are aggressively changing their holdings based on the firm-specific information. Our main interest is on the coefficient of

$DummyIndetfown \times SUE_{i,t}$. This coefficient captures the impact of the industry ETF membership on the relationship between hedge funds' abnormal holdings and the earnings surprise.

[Insert Table 4 here]

Table 4 reports our regression result. In the full sample analysis, we do not find that industry ETF membership affects the relationship between $HfAhdngRatio_{i,t}$ and $SUE_{i,t}$. But in the subsample analysis, when firms report positive earnings surprises, we find that hedge funds more aggressively increase holdings on firms with the industry ETF membership. Conversely, we find that membership has no significant impact when firms report negative earnings surprises.

The one-sided impact shows that the industry ETF influences hedge funds only when there is positive information. This asymmetry could be due to the higher costs of shorting individual stocks, making a "long the ETF/short the stock" strategy infeasible when there is negative information. Our results here are consistent with our previous finding that the industry ETF improves informational efficiency when the earnings surprise is positive. Added to our results on hedge funds abnormal holdings, we can provide evidence on the channel of this improvement: specifically, industry ETFs encourage hedge fund trading on the firm-specific information. This leads to more information being impounded into the market.

Could this effect be better explained by the trading behavior of other institutional investors? A simple placebo test is to study abnormal holdings by mutual funds. Since mutual funds are not able to short, they are unlikely to apply the "*long-the-underlying*" and "*short-the-industry-ETF*" strategy. Thus, we hypothesis that the industry ETF shall not exert any impact on mutual funds abnormal holdings regarding the earnings surprise.

[Insert Table 5 here]

We ran the same regression as in equation (4) replacing $HfAhdngRatio_{i,t}$ with $MfAhdngRatio_{i,t}$, the measure of mutual funds abnormal holdings. We construct this measure similar to the hedge funds measure, and report the summary statistics in Panel C Table 1. In Table 5, we see that industry ETF membership reduces the relationship between $MfAhdngRatio_{i,t}$ and $SUE_{i,t}$ in the full sample analysis. Most notably, it does not have any impact on the relationship when firms report positive earnings surprises. In contrast to the insignificant impact on mutual fund abnormal holdings, our previous result shows that the industry ETF increases the aggressiveness of hedge funds trading (Table 4). The sheer difference highlights the channel whereby the industry ETF improves pricing efficiency — it encourages informed trading through facilitating the long/short strategy for informed investors.

3. Predictable returns and short interest in industry ETFs

In Section 2, we showed that industry ETFs are more likely to experience extreme short interest than other ETFs, and we provide a hedging hypothesis ("*long-the-underlying*" and "*short-the-industry-ETF*") to explain this phenomenon. We show that one implication of this explanation is that industry ETFs enhance the informational efficiency of the market. That said, another implication of this hedging hypothesis is that extreme short interest should create a temporary price impact in industry ETFs. Extreme short interest creating selling pressure dampens the contemporaneous price. As the market gradually digests the shock, and realizes no industry-wide fundamental has changed, the price recovers. The price recovery implies the short interest ratio of an industry ETF, especially the change in that short interest ratio, should positively predict its future return. In contrast, the speculation hypothesis, which also explains extreme short interest, offers the different prediction. Based on the speculation hypothesis, extreme short interest reflecting the bearish speculation on industry should negatively predict the return of the corresponding industry ETF.

To test the relation between the short interest ratio and return in industry ETFs, we adopt the Fama and MacBeth (1973) approach. Every month, we regress each industry ETF's return against the change of its short interest ratio to estimate the cross-sectional correlation (the regression coefficient) between these two variables. We then calculate the time series average of these regression coefficients, and test for the significance based on the time series standard error adjusted by Newey-West with one lag. We also include an augmented regression controlling for the characteristics of the underlying firms of the industry ETF. We start our sample from 2005 due to the scarcity of industry ETFs in earlier periods. In addition, we exclude 2007 and 2008 to filter out the unusual period because of the financial crisis. The Fama-MacBeth result is reported in Table 6.

[Insert Table 6 here]

We find that the change in the short interest ratio (ΔSIR) *positively* predicts the future return for an industry ETF. To the contrary, we find the opposite for the underlying firms in the industry ETF. That is, higher short interest in a constituent firm *negatively* predicts its future return. This latter member firm-level result is consistent with past studies (see Rapach, Ringgenberg, and Zhou (2016)). Short interest on industry ETFs behaves differently from short interest on constituent firms. Our Fama-MacBeth regression result is consistent with the hedging hypothesis that extreme short interest reflects the long/short strategy from informed investors rather than the bearish speculation on the industry.

We also construct a long-short portfolio to further test the predictability of short interest on the industry ETF. We sort industry ETFs into deciles based on their ΔSIR every month. After that, we long the ETF in the highest decile, and short the ETF in the lowest decile. We evaluate the return of this long-short portfolio based on the excess return, the CAPM alpha, the Fama-French alpha, and the Fama-French-Carhart alpha. Standard errors are Newy-West adjusted with one lag. We report our long-short portfolio results in Table 7.

[Insert Table 7 here]

Our long-short portfolio generates a monthly alpha around 30 basis points, and it is significant at the 5% level. We apply the similar long-short portfolio on stocks, which are members of industry ETFs. In contrast to our findings on industry ETFs, we find the monthly alpha is around negative 30 to 40 basis points, and it is significant at the 1% level. Our long-short portfolio provides consistent evidence with the Fama-MacBeth result, i.e., the high short interest ratio on an industry ETF positively predicts the ETF's future return. This result is consistent with the implication of the hedging hypothesis. Extreme short interest reflects informed investors' hedging needs creating a temporary shock, which leads to the future price recovery.

4. Extensions and generalizations

Our analysis thus far shows strong support for the use of industry ETFs as hedging vehicles for informed traders. Our results also suggest an important role played by hedge funds in this process. What is always important to consider, however, is whether other evidence can be brought to bear to strengthen or refute our arguments. In this section, we consider two extensions to our analysis. First, we clarify the linkage of our results to informed trading by investigating how the correlation between the ETF and the underlying stock affects our "*long the stock/short the ETF*" strategy. Second, we examine more carefully whether hedge funds are acting as informed traders or whether their use of industry ETFs reflects other features particular to the unique structure of ETFs.

4.1. Correlations and industry ETF hedging

Our conjectured strategy of "*long-the-underlying*" and "*short-the-industry-ETF*" is intended to facilitate the trading of informed investors. Yet, this strategy should not work equally well across all stocks in an ETF. This is because trading on stocks with extremely high or low industry risk exposure will not benefit from establishing the hedge position. If a stock has high industry risk exposure, then it co-moves with the industry return. Consequently, the return to the stock and the ETF will be the same, and establishing a short position in the ETF is at cross purposes with the goal of profiting from information. On the other hand, if a stock has low industry risk exposure, then there is little reason to hedge with the industry ETF as the hedge will be ineffective. Hence, our strategy should not be working among these stocks either.

To test for these effects, we first compute the industry risk exposure for our sample stocks by regressing the stock-level daily return on the return of the stock-associated industry ETF. We then average the adjusted R^2 for each stock across all ETFs that include that stock. This allows us to sort the average adjusted R^2 for our sample stocks and pick the top and bottom 15% based on their industry risk exposure. Using this sample of high and low exposure stocks, we then redo our analysis on CAR and hedge funds abnormal holdings.[17] If our conjecture is correct that the ETF short position is used to hedge informed trading risk, then we should not find significant effects in this sub-sample.

[Insert Table 8 here]

[Insert Table 9 here]

Table 8 and 9 report the analysis results. Consistent with our conjecture, we do not find industry ETF membership significantly reduces the return response to the earnings surprise for stocks with extremely high or low industry risk exposures. In Table 9, we also fail to find for those stocks that the industry ETF membership leads to more aggressive hedge funds trading before the positive earnings surprise.

[Insert Table 10 here]

[Insert Table 11 here]

[17] We repeated the analysis with different cut-offs such as the top and bottom 10%, 20%, or 25%. Our results are largely consistent.

To complete the test for our conjecture, we also report the results for our sample stocks with the medium industry risk exposure, i.e., stocks with adjusted R^2 between the top and bottom 15% of our sample. Table 10 and 11 report these results. Here, we find the industry ETF membership significantly reduces the return response to the earnings surprise, especially with the positive earnings surprise. In Table 11, we find that hedge funds trade more aggressively for stocks with the industry ETF membership before positive earnings surprises. Both results are consistent with our earlier findings and underscore the important hedging role played by the ETF in facilitating informed trading.

4.2. Hedging risk or arbitrage of industry ETFs

One of the important innovations underlying the ETF structure is the creation and redemption process. This daily settling-up ensures that the price of the ETF and the value of the underlying constituent stocks stays within tight bounds. If the ETF becomes overpriced, market arbitrageurs (known as Authorized Participants) will sell the ETF, buy the underlying stocks, and at the end of the day present the bundle of stocks to the ETF provider for a new ETF, thereby settling their short position. If the stocks are overpriced, the process works in reverse.[18] Hedge funds can be participants in this process, and in the case of industry ETFs this strategy implies arbitraging through loading on industry risks. This behavior raises the concern that the positive correlation between the ratio of hedge funds positive abnormal holdings and short interest in ETF is due solely to arbitrage activity and not to information-based trading. If this were the case, then we would expect to find that short interest which also relates to industry risks should be positively associated with the aggregate hedge funds' trading.

[18] Note that this process is inherently symmetric as arbitraging divergences requires going long in stocks if the ETF is overpriced and short in stocks if the ETF is underpriced. As there is no reason to believe ETFs are systematically mispriced in one direction, this symmetry seems unlikely to explain the asymmetric effects we identify.

To address this concern, we compute for each industry ETF the ratio of its constituent stocks that have positive abnormal hedge fund holdings *(%HfPosAH*) in the past quarter. Specifically, for industry ETF i (we assume this industry ETF has K member stocks), we denote its constituent stock as j, where $j = 1,2, \dots, K$. For one specific quarter, if stock j in industry ETF i has positive abnormal hedge fund holdings *(%HfPosAH*), a dummy variable equals 1. Otherwise, a dummy variable equals 0. We then calculate the percentage of member stocks having dummy variable of 1. The *%HfPosAH* variable is thus an ETF level measure, and it captures on aggregate how aggressively hedge funds are betting on the industry risk of the underlying industry tracked by the ETF. Table 12 Panel A gives the summary statistics of the measure.

[Insert Table 12 here]

To explore the relationship between this ratio of hedge funds positive abnormal holdings and the ETF's lagged short interest (SIR_{t-1}), we run the following regression:

$$\%HfPosAH_{i,t} = \alpha_i + \alpha_t + \delta SIR_{i,t-1} + controls + \epsilon_{i,t}. \tag{5}$$

Panel B of Table 12 reports the regression result. We find that the relation between *%HfPosAH* and SIR_{t-1} is insignificant. This suggests that the lagged short interest does not predict the industry-wide hedge funds positions. This is consistent with our hypothesis of a "*long-the-underlying*" and "*short-the-industry-ETF*" strategy being used to isolate the position on the firm-specific risk.[19] Because hedge funds bet on firm-specific risks of the ETF members, the aggregate hedge funds' trading (betting on firm-specific risks) in underlying stocks is not related to industry risks, and thus is not related to ETF short interest.

[19]The stock return consists of three components, market risks, industry-level risks, and idiosyncratic firm-specific risks (see Campbell et al. (2001)). While market risks can be hedged with index futures, disaggregating the latter two is more difficult (due to the lack of instruments or costly to implement). The "*long-the-underlying*" and "*short-the-industry-ETF*" enables investors to deal with this difficulty. That is, when hedge funds have information about firm-specific components, the aforementioned long/short strategy facilitates investors to isolate their positions on firm-specific risks. Here, the short position on the industry ETF serves as the hedge of industry risks.

5. Conclusions

Can industry ETFs facilitate informed trading and enhance the efficiency of the market? Our results show that they can by facilitating the hedging of industry risk for informed investors. We demonstrate that because of this hedging role increased short interest in industry ETFs is a bullish, not bearish, signal of future performance. Using earnings announcement events, we show that industry ETF short interest predicts more positive earnings surprises among its underlying stocks. We also find that it reduces the market reaction to positive earnings surprises and leads to *ex-ante* more trading by hedge funds. These effects are consistent with industry ETFs increasing informed trading in individual stocks, thereby making the market more informationally efficient. We also showed that the change in industry short interest predicts the future return for the industry ETF, a result we ascribe to the hedging-based use of the ETF inducing a temporary price impact that reverses when new industry information does not materialize. Overall, industry ETFs appear to be a valuable innovation in the market.

One aspect of our results that we find particularly intriguing is the asymmetry of the effects: these positive effects on the market arise only with positive news about firms and not negative news. We believe this reflects another aspect of this financial innovation as industry ETFs reduce the transactions cost of shorting, making the "*long the stock/short the ETF*" strategy feasible. No similar innovation exists to reduce the shorting costs of individual stocks, but perhaps future financial innovation can address this problem.

References

Allen, Franklin, and Douglas Gale, 1994, *Financial Innovation and Risk Sharing* (MIT Press).

Battalio, Robert, and Paul Schultz, 2011, Regulatory uncertainty and market liquidity: The 2008 short sale ban's impact on equity option markets, *Journal of Finance* 66, 2013–2053.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2011, Hedge fund stock trading in the financial crisis of 2007–2009, *Review of Financial Studies* 25, 1–54.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2014, Do ETFs increase volatility? *Working Paper*.

Bhattacharya, Ayan, and Maureen O'Hara, 2017, Can ETFs increase market fragility? Effect of information linkages in ETF markets, *Working Paper*.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang, 2008, Which shorts are informed? *Journal of Finance* 63, 491–527. Brunnermeier, Markus K., and Stefan Nagel, 2005, Hedge funds and the technology bubble, *Journal of Finance* 59, 2013–2040.

Brunnermeier, Markus K., and Stefan Nagel, 2005, Hedge funds and the technology bubble, *Journal of Finance* 59, 2013–2040.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, 2001, Have individual stocks become more volatile? An Exploration of idiosyncratic risk" Journal of Finance, 56, 1-43.

Chen, Yong, Zhi Da, and Dayong Huang, 2016, Arbitrage trading: The long and the short of it, *Working Paper*, 1–56.

Chen, Zhiwu, 1995, Financial innovation and arbitrage pricing in frictional economies, *Journal of Economic Theory* 65, 117–135.

Da, Zhi, and Sophie Shive, 2015, When the bellwether dances to noise: Evidence from exchange-traded funds, *Working Paper*.

Desai, Hemang, K. Ramesh, S. Ramu Thiagarajan, and Bala V. Balachandran, 2002, An investigation of the informational role of short interest in the Nasdaq market, *Journal of Finance* 57, 2263–2287.

Diether, Karl B., Kuan-Hui Lee, and Ingrid M. Werner, 2009, Short-sale strategies and return predictability, *Review of Financial Studies* 22, 575–607.

Dow, James, 1998, Arbitrage, hedging, and financial innovation, *Review of Financial Studies* 11, 739–755.

Duffie, Darrell, and Rohit Rahi, 1995, Financial market innovation and security design: An introduction, *Journal of Economic Theory* 65, 1–42.

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

Glosten, Lawrence, Suresh Nallareddy, and Yuan Zou, 2017, ETF activity and informational efficiency of underlying securities, *Working Paper*.

Griffin, John M., and Jin Xu, 2009, How smart are the smart guys? A unique view from hedge fund stock holdings, *Review of Financial Studies* 22, 2531–2570.

Grundy, Bruce D., Bryan Lim, and Patrick Verwijmeren, 2012, Do option markets undo restrictions on short sales? Evidence from the 2008 short-sale ban, *Journal of Financial Economics* 106, 331–348.

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

Jiang, Wenxi, 2017, Leveraged speculators and asset prices, *Working Paper*, 1–59.

Li, Frank Weikai, and Qifei Zhu, 2017, Short selling ETFs, *Working Paper*, 1–68.

Livnat, Joshua, and Richard R. Mendenhall, 2006, Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts, *Journal of Accounting Research* 44, 177–205.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457–3489.

Madhavan, Ananth, and Aleksander Sobczyk, 2015, Price dynamics and liquidity of exchange-traded funds, *Working Paper*.

Ofek, Eli, Matthew Richardson, and Robert F. Whitelaw, 2004, Limited arbitrage and short sales restrictions: Evidence from the options markets, *Journal of Financial Economics* 74, 305–342.

Rapach, David E., Matthew C. Ringgenberg, and Guofu Zhou, 2016, Short interest and aggregate stock returns, *Journal of Financial Economics* 121, 46–65.

Simsek, Alp, 2013a, Financial innovation and portfolio risks, *American Economic Review* 103, 398–401.

Simsek, Alp, 2013b, Speculation and risk sharing with new financial assets, *Quarterly Journal of Economics* 128, 1365–1396.

Wermers, Russ, and Jinming Xue, 2015, Intraday ETF trading and the volatility of the underlying, *Working Paper*.

Table 1: Summary statistics for the sample

Panel A and B report the summary statistics on quarter short interest ratio (*SIR*), price, volume, and total net asset value (*TNA*) for industry and non-industry ETFs, respectively. The quarter measure is constructed by taking the average of monthly observations. Panel C reports the summary statistics for firms in our earnings announcement sample. *DummyIndetfown* is the dummy variable which equals to 1 if the firm is a member of an industry ETF. *CAR* is the -1 to +1 cumulative abnormal daily return around the earnings announcement date based on the Fama-French three factor model. *SUE* is the standardized earnings surprise computed from a rolling seasonal random walk model. *HfAhdngRatio* and *MfAhdngRatio* are the abnormal holdings from hedge funds and mutual funds, respectively. Both holdings are standardized by shares outstanding. *log(MktCap)* is the log transformed market capitalization. *BM* is the book-to-market ratio where the book value is measured as the preceding fiscal year, and market value is measured as of the end of that calendar year. *TR* is the turnover measured as the average of the daily ratios of volume over shares outstanding from -40 to -11 of each announcement. *MOM* is the cumulative raw return over the six-month period ending one month prior to the announcement month. *EarnPerst* is the earnings persistence as of the first-order auto-regressive coefficient of quarterly earnings over the past four years. *NumEst* is the number of analysts. Panel D reports the summary statistics on the ratio of positive *SUE* in an ETF over our sample period. The positive *SUE* is defined as the *SUE* exceeding the 75 percentile of all *SUE*s in the sample.

Panel A: Industry ETFs

	Mean	**Std.**	**5%**	**25%**	**50%**	**75%**	**95%**
SIR	0.118	0.211	0.001	0.007	0.026	0.117	**0.607**
Price	53.842	32.056	18.104	30.167	48.465	68.034	110.976
Volume (in shares)	422847.684	2111222.625	764.333	5382.917	19368.167	75731.833	1963009.083
TNA (in $ millions)	1051.615	2156.886	11.417	91.817	331.200	918.467	5295.167

Panel B: Non-industry ETFs

	Mean	**Std.**	**5%**	**25%**	**50%**	**75%**	**95%**
SIR	0.041	0.118	0.000	0.003	0.008	0.021	**0.191**
Price	56.290	35.659	15.884	28.185	48.853	75.056	123.672
Volume (in shares)	313927.645	2855881.357	538.017	2304.917	7388.500	32629.167	380683.450
TNA (in $ millions)	2309.432	9704.713	9.467	50.733	179.333	923.600	10578.307

Panel C: The earnings announcement sample

	Mean	Std.	5%	25%	50%	75%	95%
DummyIndetfown	0.414	0.493	0.000	0.000	0.000	1.000	1.000
CAR	0.001	0.096	-0.142	-0.038	0.000	0.041	0.145
SUE	-0.002	0.059	-0.072	-0.006	0.001	0.007	0.060
HfAhdngRatio	0.198	2.044	-2.985	-0.304	0.000	0.546	4.094
MfAhdngRatio	0.600	5.826	-8.819	-1.834	0.155	2.875	11.148
log(MktCap)	19.785	1.985	16.778	18.305	19.649	21.098	23.301
BM	0.668	0.573	0.102	0.296	0.526	0.853	1.718
TR	0.006	0.006	0.000	0.002	0.004	0.008	0.019
MOM	0.075	0.373	-0.468	-0.140	0.039	0.226	0.750
EarnPerst	0.264	0.363	-0.263	-0.011	0.218	0.528	0.903
NumEst	4.020	5.400	0.000	0.000	2.000	6.000	16.000

Panel D: The ratio of positive SUE in an ETF

	Mean	Std.	5%	25%	50%	75%	95%
Ind. ETFs	0.164	0.111	0.038	0.090	0.137	0.205	0.394
Non-ind. ETFs	0.153	0.084	0.049	0.098	0.138	0.190	0.313

Table 2: Regress the positive earnings ratio on the lagged short interest at the ETF level

Table 2 reports the result of regressing the positive earnings ratio ($PosSUE_{i,t}$) on the lagged short interest ratio ($SIR_{i,t-1}$), i.e.,

$$PosSUE_{i,t} = \alpha_i + \alpha_t + \beta_1 SIR_{i,t-1} + controls + \epsilon_{i,t}, \quad \text{(A1)}$$

$$PosSUE_{i,t} = \alpha_i + \alpha_t + \beta_1 SIR_{i,t-1} + \beta_2 DummyIndetf \times SIR_{i,t-1} + controls + \epsilon_{i,t}. \quad \text{(A2)}$$

We run the first regression model on industry and non-industry ETFs, respectively. And we use the second regression model to estimate the difference in the predicting power of the short interest ratio to the positive earnings ratio between industry and non-industry ETFs. The difference is captured by β_2 (the coefficient on $DummyIndetf \times SIR_{i,t-1}$). In our controls, we include log total net asset value, and year, quarter, and ETF fixed effects. All standard errors are clustered by ETFs and quarters. Panel A reports the result on full sample, and Panel B reports the result on full sample excluding year 2007 and 2008. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Panel A: Regression result on the full sample

Dependent variable: PosSUE_(i,t)	**Ind. ETFs**	**Non-ind. ETFs**	**All**
SIR_(i,t-1)	0.0324*	-0.0235*	-0.0253*
	(1.6687)	(-1.6823)	(-1.7972)
DummyIndetf x SIR_(i,t-1)	-	-	**0.0581****
	-	-	(2.481)
log(TNA)	0.0067**	0.0024**	0.0037***
	(2.416)	(2.2304)	(3.0643)
Year F.E.	Y	Y	Y
Quarter F.E.	Y	Y	Y
ETF F.E.	Y	Y	Y
Num.Obs.	4079	9413	13492
R-squared	0.4868	0.6592	0.5855

Panel B: Regression result on the full sample excluding 2007 and 2008

Dependent variable: PosSUE_(i,t)	**Ind. ETFs**	**Non-ind. ETFs**	**All**
SIR_(i,t-1)	0.0436*	-0.0124	-0.0174
	(1.8814)	(-1.0757)	(-1.4781)
DummyIndetf x SIR_(i,t-1)	-	-	**0.0625****
	-	-	(2.3364)
log(TNA)	0.0099***	0.0043***	0.0058***
	(2.7909)	(3.2452)	(3.8403)
Year F.E.	Y	Y	Y
Quarter F.E.	Y	Y	Y
ETF F.E.	Y	Y	Y
Num.Obs.	3580	8244	11824
R-squared	0.5003	0.6817	0.6054

Table 3: The regression result on the market reaction to the earnings surprise

Table 3 reports the regression result on the following model,

$$CAR_{i,t} = \alpha_i + \alpha_t + \theta_1 SUE_{i,t} + \theta_2 DummyIndetfown + \theta_3 DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}. \quad (A3)$$

$CAR_{i,t}$ is the -1 to +1 cumulative abnormal daily return around the earnings announcement date based on the Fama-French three factor model. $SUE_{i,t}$ is the standardized earnings surprise computed from a rolling seasonal random walk model. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an industry ETF. $DummyIndetfown \times SUE_{i,t}$ is the interaction term. In our control variables, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts. In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008. Furthermore, we split our earnings announcements sample into the "Negative SUE" and "Positive SUE" group based on the earnings surprise, $SUE_{i,t}$. The "Negative SUE" group consists of the bottom 25% $SUE_{i,t}$ in our sample, while the "Positive SUE" group consists of the top 25% $SUE_{i,t}$ in our sample. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Dependent variable: CAR_(i,t)	**All**	**Negative SUE**	**Positive SUE**
SUE_(i,t)	0.2002***	0.0944***	0.0043
	(30.8998)	(9.6612)	(0.3632)
DummyIndetfown	0.007***	0.0045***	0.0153***
	(9.6669)	(3.0769)	(9.765)
DummyIndetfown x SUE_(i,t)	**-0.0551*** **	-0.014	**-0.0464** **
	(-5.0421)	(-0.8049)	(-2.4792)
log(MktCap)	-0.0014***	0.0011***	-0.0072***
	(-7.3373)	(2.9473)	(-16.2583)
BM	0.0032***	0.0036***	0.0039***
	(7.3859)	(4.8421)	(4.9137)
TR	-0.566***	-0.5242***	-0.8457***
	(-10.0016)	(-5.4694)	(-8.2553)
MOM	-0.0012	-0.0071***	-0.0021
	(-1.3935)	(-4.9633)	(-1.5745)
EarnPerst	0.0	-0.0016	0.0018
	(0.0429)	(-1.3259)	(1.3534)
NumEst	0.0004***	0.0007***	0.0007***
	(7.5835)	(5.7893)	(6.0641)
Industry F.E.	Y	Y	Y
Month F.E.	Y	Y	Y
Year F.E.	Y	Y	Y
Num.Obs.	291599	72715	72922
R-squared	0.0163	0.0116	0.0155

Table 4: Hedge funds abnormal holdings and the earnings surprise

Table 4 reports the regression result on the following model,

$$HfAhdngRatio_{i,t}, \quad \text{(A4)}$$
$$= \alpha_i + \alpha_t + \theta_1^H SUE_{i,t} + \theta_2^H DummyIndetfown$$
$$+ \theta_3^H DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}.$$

$HfAhdngRatio_{i,t}$ is the abnormal holdings by hedge funds standardized by total shares outstanding. The abnormal holdings is estimated as the difference between the current quarter holdings and the moving average of holdings in past four quarters. $SUE_{i,t}$ is the standardized earnings surprise computed from a rolling seasonal random walk model. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an industry ETF. $DummyIndetfown \times SUE_{i,t}$ is the interaction term. In our control variables, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts. In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008. Furthermore, we split our earnings announcements sample into the "Negative SUE" and "Positive SUE" group based on the earnings surprise, $SUE_{i,t}$. The "Negative SUE" group consists of the bottom 25% $SUE_{i,t}$ in our sample, while the "Positive SUE" group consists of the top 25% $SUE_{i,t}$ in our sample. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Dependent variable: HfAhdngRatio_(i,t)	**All**	**Negative SUE**	**Positive SUE**
SUE_(i,t)	0.1345	0.0428	0.2352
	(1.5702)	(0.3039)	(1.3629)
DummyIndetfown	-0.0363*	0.0164	-0.1041***
	(-1.9484)	(0.4293)	(-2.8134)
DummyIndetfown x *SUE_(i,t)*	0.0735	-0.0181	**0.931****
	(0.3437)	(-0.0554)	(2.1608)
log(MktCap)	-0.0122***	0.0108	0.0226***
	(-3.1918)	(1.3304)	(2.9731)
BM	0.0197*	-0.0169	-0.007
	(1.8528)	(-0.9672)	(-0.4304)
TR	32.3415***	23.336***	31.3782***
	(23.6757)	(9.7707)	(13.2446)
MOM	0.034*	0.1052***	0.0972***
	(1.9494)	(3.5782)	(4.0837)
EarnPerst	-0.0672***	0.0074	-0.0444
	(-4.7204)	(0.2544)	(-1.5082)
NumEst	-0.007***	-0.0035	-0.0124***
	(-5.2105)	(-1.1806)	(-4.0509)
Industry F.E.	Y	Y	Y
Month F.E.	Y	Y	Y
Year F.E.	Y	Y	Y
Num.Obs.	291620	72722	72932
R-squared	0.035	0.0397	0.0342

Table 5: Mutual funds abnormal holdings and the earnings surprise

Table 5 reports the regression result on the following model,

$$MfAhdngRatio_{i,t}, \quad (A5)$$
$$= \alpha_i + \alpha_t + \theta_1^M SUE_{i,t} + \theta_2^M DummyIndetfown$$
$$+ \theta_3^M DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}.$$

$MfAhdngRatio_{i,t}$ is the abnormal holdings by mutual funds standardized by total shares outstanding. The abnormal holdings is estimated as the difference between the current quarter holdings and the moving average of holdings in past four quarters. $SUE_{i,t}$ is the standardized earnings surprise computed from a rolling seasonal random walk model. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an industry ETF. $DummyIndetfown \times SUE_{i,t}$ is the interaction term. In our control variables, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts. In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008. Furthermore, we split our earnings announcements sample into the "Negative SUE" and "Positive SUE" group based on the earnings surprise, $SUE_{i,t}$. The "Negative SUE" group consists of the bottom 25% $SUE_{i,t}$ in our sample, while the "Positive SUE" group consists of the top 25% $SUE_{i,t}$ in our sample. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Dependent variable: MfAhdngRatio_(i,t)	**All**	**Negative SUE**	**Positive SUE**
SUE_(i,t)	2.5139***	4.0446***	-4.2241***
	(10.6551)	(10.0637)	(-10.0624)
DummyIndetfown	-0.9834***	-0.5686***	-1.1736***
	(-14.9465)	(-5.2105)	(-11.6335)
DummyIndetfown x SUE_(i,t)	-0.6845	0.6863	0.8297
	(-1.1713)	(0.7723)	(0.815)
log(MktCap)	0.3333***	0.3589***	0.4869***
	(19.7558)	(13.521)	(19.4775)
BM	-0.5808***	-0.3701***	-0.381***
	(-16.2149)	(-7.523)	(-8.6381)
TR	71.5907***	-15.8043**	82.3471***
	(16.9328)	(-2.442)	(12.9173)
MOM	2.4397***	2.38***	1.5836***
	(34.256)	(25.9914)	(20.9999)
EarnPerst	0.1002**	-0.0479	0.1877**
	(2.1033)	(-0.5969)	(2.2023)
NumEst	-0.0776***	-0.0396***	-0.0844***
	(-16.9889)	(-4.6167)	(-9.7255)
Industry F.E.	Y	Y	Y
Month F.E.	Y	Y	Y
Year F.E.	Y	Y	Y
Num.Obs.	291620	72722	72932
R-squared	0.0896	0.0694	0.0898

Table 6: Fama-MacBeth regression of returns on short interest ratios

Table 6 reports the time-series averages of slope coefficients from Fama and MacBeth (1973) cross-sectional regressions on returns, Ret_{t+1}, and changes in the short interest ratio, ΔSIR_t, for industry ETFs and their member firms, respectively. t-statistic reported in the parenthesis is calculated using the average slope coefficient divided by its time-series standard error adjusted by Newey-West with one lag. For each industry ETF, we average the member firms' characteristics, and use the average as a control in our regression. We also repeat the Fama-MacBeth regression on industry ETFs' member firms, and report in the last two columns. For the firm level regression, the control variable corresponds to each firm's own characteristics. The sample period of our analysis is reported in the last row of the table. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%.

Dependent variable: Ret_(t+1)	**Industry ETFs**		**Member firms**	
ΔSIR_t	0.019**	0.024***	-0.064**	-0.080***
	(2.27)	(2.65)	(-2.32)	(-2.83)
Ret from t-12 to t-1	-	-0.007	-	0.000
		(-0.59)		(-0.15)
Ret in month t	-	0.003	-	-0.014**
		(0.09)		(-2.13)
BM	-	-0.001	-	0.002**
		(-0.09)		(2.12)
Operating profit	-	0.001	-	0.000
		(0.18)		(0.67)
Market capitalization	-	0.000	-	0.000
		(-0.91)		(-1.56)
Asset growth	-	0.004	-	-0.002
		(0.25)		(-1.45)
Investment growth	-	0.005	-	0.000
		(1.32)		(-1.36)
Gross profitability	-	-0.002	-	0.004*
		(-0.18)		(1.70)
Net stock issuance	-	-0.009	-	0.000
		(-0.39)		(0.09)
Accruals	-	0.000	-	-0.002**
		(0.00)		(-2.08)
Net operating assets	-	-0.007	-	-0.002
		(-0.91)		(-0.83)
Intercept	0.013***	0.010***	0.009**	0.007*
	(3.55)	(3.25)	(2.46)	(1.71)
Sample period	2005.01 - 2006.12, 2009.01 - 2016.11		1999.03 - 2016.11	

Table 7: Long-short portfolio sorts on ΔSIR

Table 7 reports the average monthly excess returns, CAPM alpha, Fama and French 3-factor alpha, and Fama-French-Carhart 4-factor alpha for each of the 10 decile portfolios and the High-Low portfolio based on ΔSIR. At the end of each month, all industry ETFs or member firms are sorted into deciles based on ΔSIR in that month. Then, we track the equal-weighted portfolio returns over the next month. Panel A reports results for industry ETFs. Holding periods in Panel A is from January 2005 to December 2016, excluding sample in year 2007 and 2008. Panel B reports results for member firms. Member firms with prices below $5 a share or are in the bottom NYSE size decile are excluded from the sample. Holding period in Panel B is from April 1999 to December 2016. Standard errors are Newey-West adjusted with one lag.

Panel A: Industry ETFs

	Excess returns		**CAPM alpha**		**3-factor alpha**		**4-factor alpha**	
Decile	coef.	t-stat.	coef.	t-stat.	coef.	t-stat.	coef.	t-stat.
1 (low)	1.16	2.79	-0.27	-2.29	-0.23	-2.35	-0.20	-2.12
2	1.25	3.06	-0.15	-1.44	-0.12	-1.22	-0.09	-0.91
3	1.37	3.69	-0.17	-1.56	-0.14	-1.33	-0.13	-1.23
4	1.45	4.01	-0.01	-0.08	0.04	0.36	0.04	0.36
5	1.36	3.82	0.07	0.69	0.12	1.22	0.13	1.35
6	1.42	3.79	-0.11	-0.67	-0.04	-0.30	-0.05	-0.33
7	1.36	3.63	0.01	0.13	0.05	0.48	0.06	0.57
8	1.15	2.99	-0.06	-0.55	-0.03	-0.27	-0.01	-0.09
9	1.46	3.65	-0.14	-1.15	-0.10	-0.85	-0.10	-0.83
10 (high)	1.53	3.77	0.02	0.16	0.05	0.44	0.08	0.65
10-1	**0.28**	**2.00**	**0.29**	**2.00**	**0.28**	**1.97**	**0.28**	**1.98**

Panel B: Member firms

	Excess returns		**CAPM alpha**		**3-factor alpha**		**4-factor alpha**	
Decile	coef.	t-stat.	coef.	t-stat.	coef.	t-stat.	coef.	t-stat.
1 (low)	0.81	1.93	0.32	1.55	0.01	0.09	0.05	0.32
2	0.94	2.47	0.49	2.67	0.22	1.67	0.25	1.92
3	0.90	2.53	0.49	2.74	0.24	1.80	0.26	1.93
4	0.98	2.85	0.58	3.32	0.33	2.72	0.34	2.77
5	0.90	2.60	0.52	3.03	0.28	2.29	0.29	2.39
6	1.02	3.04	0.64	3.65	0.37	3.18	0.38	3.24
7	0.94	2.63	0.54	3.06	0.27	2.38	0.29	2.53
8	0.87	2.35	0.44	2.45	0.21	1.52	0.23	1.65
9	0.62	1.59	0.18	0.93	-0.10	-0.71	-0.06	-0.45
10 (high)	0.48	1.08	-0.04	-0.19	-0.35	-2.20	-0.30	-1.95
10-1	**-0.33**	**-3.13**	**-0.36**	**-3.37**	**-0.36**	**-3.31**	**-0.35**	**-3.10**

Table 8: The regression result on the market reaction to the earnings surprise

Table 8 reports the regression result on the following model with stocks belong to the top and bottom 15% sorting on the industry risk exposure (the average adjusted R^2 when regressing stock returns on industry ETF returns),

$$CAR_{i,t} = \alpha_i + \alpha_t + \theta_1 SUE_{i,t} + \theta_2 DummyIndetfown + \theta_3 DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}. \quad \text{(A6)}$$

$CAR_{i,t}$ is the -1 to +1 cumulative abnormal daily return around the earnings announcement date based on the Fama-French three factor model. $SUE_{i,t}$ is the standardized earnings surprise computed from a rolling seasonal random walk model. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an industry ETF. $DummyIndetfown \times SUE_{i,t}$ is the interaction term. In our control variables, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts. In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008. Furthermore, we split our earnings announcements sample into the "Negative SUE" and "Positive SUE" group based on the earnings surprise, $SUE_{i,t}$. The "Negative SUE" group consists of the bottom 25% $SUE_{i,t}$ in our sample, while the "Positive SUE" group consists of the top 25% $SUE_{i,t}$ in our sample. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Dependent variable: CAR_(i,t)	**All**	**Negative SUE**	**Positive SUE**
SUE_(i,t)	0.1488***	0.0695**	0.0082
	(6.7297)	(2.1449)	(0.2033)
DummyIndetfown	0.0039***	0.0059*	0.0049
	(2.7318)	(1.8934)	(1.3376)
DummyIndetfown x SUE_(i,t)	**-0.002**	**0.0691***	**-0.0174**
	(-0.0751)	(1.6913)	(-0.3805)
log(MktCap)	-0.0004	0.0009	-0.0032***
	(-1.1)	(1.1456)	(-3.9537)
BM	0.005***	0.0097***	0.0036*
	(4.7764)	(4.8867)	(1.7696)
TR	-0.1763*	-0.0443	-0.2945*
	(-1.6504)	(-0.263)	(-1.7148)
MOM	-0.0032*	-0.0071**	-0.005
	(-1.8633)	(-2.4448)	(-1.6382)
EarnPerst	0.0015	0.0017	0.0017
	(1.4832)	(0.6914)	(0.6916)
NumEst	0.0002*	0.0003*	0.0002
	(1.8795)	(1.6926)	(1.0253)
Industry F.E.	Y	Y	Y
Month F.E.	Y	Y	Y
Year F.E.	Y	Y	Y
Num.Obs.	53111	13250	13269
R-squared	0.0118	0.0169	0.0078

Table 9: Hedge funds abnormal holdings and the earnings surprise

Table 9 reports the regression result on the following model with stocks belong to the top and bottom 15% sorting on the industry risk exposure (the average adjusted R^2 when regressing stock returns on industry ETF returns),

$$HfAhdngRatio_{i,t}, \quad \text{(A7)}$$

$$= \alpha_i + \alpha_t + \theta_1^H SUE_{i,t} + \theta_2^H DummyIndetfown$$

$$+ \theta_3^H DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}.$$

$HfAhdngRatio_{i,t}$ is the abnormal holdings by hedge funds standardized by total shares outstanding. The abnormal holdings is estimated as the difference between the current quarter holdings and the moving average of holdings in past four quarters. $SUE_{i,t}$ is the standardized earnings surprise computed from a rolling seasonal random walk model. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an industry ETF. $DummyIndetfown \times SUE_{i,t}$ is the interaction term. In our control variables, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts. In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008. Furthermore, we split our earnings announcements sample into the "Negative SUE" and "Positive SUE" group based on the earnings surprise, $SUE_{i,t}$. The "Negative SUE" group consists of the bottom 25% $SUE_{i,t}$ in our sample, while the "Positive SUE" group consists of the top 25% $SUE_{i,t}$ in our sample. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Dependent variable: HfAhdngRatio_(i,t)	**All**	**Negative SUE**	**Positive SUE**
SUE_(i,t)	0.5135	0.4047	0.8272
	(1.2521)	(0.6163)	(1.0709)
DummyIndetfown	-0.0287	-0.0457	-0.1208
	(-0.6538)	(-0.5486)	(-1.3887)
DummyIndetfown x *SUE_(i,t)*	**-0.4291**	**-0.0999**	**0.2364**
	(-0.7416)	(-0.1218)	(0.2184)
log(MktCap)	-0.0206**	0.0076	-0.0174
	(-2.2406)	(0.3843)	(-0.9766)
BM	0.0132	-0.0074	-0.0202
	(0.4666)	(-0.1349)	(-0.4104)
TR	36.9965***	37.7377***	30.5388***
	(12.2542)	(8.1556)	(5.9874)
MOM	-0.0265	0.1294	0.0613
	(-0.6183)	(1.623)	(0.9452)
EarnPerst	-0.0669**	-0.0584	-0.0291
	(-2.2588)	(-0.8583)	(-0.4514)
NumEst	-0.0056**	-0.0067	-0.0054
	(-2.5527)	(-1.4078)	(-1.2634)
Industry F.E.	Y	Y	Y
Month F.E.	Y	Y	Y
Year F.E.	Y	Y	Y
Num.Obs.	53113	13251	13270
R-squared	0.0461	0.064	0.0404

Table 10: The regression result on the market reaction to the earnings surprise

Table 10 reports the regression result on the following model with stocks between the top and bottom 15% sample of the industry risk exposure (the average adjusted R^2 when regressing stock returns on industry ETF returns),

$$CAR_{i,t} = \alpha_i + \alpha_t + \theta_1 SUE_{i,t} + \theta_2 DummyIndetfown + \theta_3 DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}. \quad \text{(A8)}$$

$CAR_{i,t}$ is the -1 to +1 cumulative abnormal daily return around the earnings announcement date based on the Fama-French three factor model. $SUE_{i,t}$ is the standardized earnings surprise computed from a rolling seasonal random walk model. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an industry ETF. $DummyIndetfown \times SUE_{i,t}$ is the interaction term. In our control variables, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts. In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008. Furthermore, we split our earnings announcements sample into the "Negative SUE" and "Positive SUE" group based on the earnings surprise, $SUE_{i,t}$. The "Negative SUE" group consists of the bottom 25% $SUE_{i,t}$ in our sample, while the "Positive SUE" group consists of the top 25% $SUE_{i,t}$ in our sample. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Dependent variable: CAR_(i,t)	**All**	**Negative SUE**	**Positive SUE**
SUE_(i,t)	0.2042***	0.0942***	0.0015
	(30.2763)	(9.2223)	(0.1192)
DummyIndetfown	0.0081***	0.0057***	0.0182***
	(9.6302)	(3.389)	(9.9366)
DummyIndetfown x SUE_(i,t)	**-0.0591*** **	**-0.0206**	**-0.0496** **
	(-4.8172)	(-1.0505)	(-2.303)
log(MktCap)	-0.0016***	0.0011**	-0.0081***
	(-7.5158)	(2.5243)	(-15.2923)
BM	0.0029***	0.0029***	0.0037***
	(6.0221)	(3.5225)	(4.3787)
TR	-0.6552***	-0.6696***	-0.9804***
	(-10.703)	(-6.038)	(-8.2912)
MOM	-0.0008	-0.0071***	-0.0012
	(-0.8655)	(-4.5292)	(-0.8172)
EarnPerst	-0.0003	-0.0022	0.0013
	(-0.5245)	(-1.62)	(0.8673)
NumEst	0.0005***	0.0008***	0.0009***
	(7.2109)	(5.4286)	(4.9705)
Industry F.E.	Y	Y	Y
Month F.E.	Y	Y	Y
Year F.E.	Y	Y	Y
Num.Obs.	238488	59463	59643
R-squared	0.0173	0.0114	0.0165

Table 11: Hedge funds abnormal holdings and the earnings surprise

Table 11 reports the regression result on the following model with stocks between the top and bottom 15% sample of the industry risk exposure (the average adjusted R^2 when regressing stock returns on industry ETF returns),

$$
\begin{aligned}
HfAhdngRatio_{i,t}, \quad & \quad (A9)\\
= \ & \alpha_i + \alpha_t + \theta_1^H SUE_{i,t} + \theta_2^H DummyIndetfown \\
+ \ & \theta_3^H DummyIndetfown \times SUE_{i,t} + controls + \epsilon_{i,t}.
\end{aligned}
$$

$HfAhdngRatio_{i,t}$ is the abnormal holdings by hedge funds standardized by total shares outstanding. The abnormal holdings is estimated as the difference between the current quarter holdings and the moving average of holdings in past four quarters. $SUE_{i,t}$ is the standardized earnings surprise computed from a rolling seasonal random walk model. $DummyIndetfown$ is a dummy variable, which equals to 1 if the firm is a constituent of an industry ETF. $DummyIndetfown \times SUE_{i,t}$ is the interaction term. In our control variables, we include the market capitalization, the book-to-market ratio, the turnover, the momentum factor, the earnings persistence, and the number of analysts. In addition, we also control for industry, month, and year fixed effects. All standard errors are clustered by firms and announcement dates. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008. Furthermore, we split our earnings announcements sample into the "Negative SUE" and "Positive SUE" group based on the earnings surprise, $SUE_{i,t}$. The "Negative SUE" group consists of the bottom 25% $SUE_{i,t}$ in our sample, while the "Positive SUE" group consists of the top 25% $SUE_{i,t}$ in our sample. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%. t-statistic is reported in the parenthesis.

Dependent variable: HfAhdngRatio_(i,t)	**All**	**Negative SUE**	**Positive SUE**
SUE_(i,t)	0.0904	0.0262	0.2201
	(1.0549)	(0.1832)	(1.2509)
DummyIndetfown	-0.0434**	0.0005	-0.0989**
	(-2.1076)	(0.0121)	(-2.3265)
DummyIndetfown x *SUE_(i,t)*	**0.1868**	**-0.0369**	**0.9283***
	(0.7998)	(-0.1)	(1.8971)
log(MktCap)	-0.0087**	0.0187**	0.0339***
	(-2.0472)	(2.0127)	(3.8534)
BM	0.0198*	-0.0168	-0.0055
	(1.7249)	(-0.891)	(-0.3203)
TR	30.9998***	20.4095***	30.6896***
	(20.7515)	(7.5831)	(11.7445)
MOM	0.0452***	0.0984***	0.1022***
	(2.5856)	(3.2087)	(4.095)
EarnPerst	-0.0678***	0.0155	-0.0391
	(-4.2454)	(0.4926)	(-1.206)
NumEst	-0.0079***	-0.0034	-0.0145***
	(-4.7335)	(-0.8544)	(-3.4837)
Industry F.E.	Y	Y	Y
Month F.E.	Y	Y	Y
Year F.E.	Y	Y	Y
Num.Obs.	238507	59470	59652
R-squared	0.0331	0.0352	0.0325

Table 12: Aggregate Hedge Fund Abnormal Holding and Lagged Short Interest

Table 12 Panel A reports the summary statistics of the positive abnormal hedge fund holdings. For each industry ETF, %HfPosAH is computed as the percentage of its members that have positive abnormal hedge fund holdings. Panel B reports the result of regressing the ratio of positive hedge funds abnormal holding of an industry ETF ($\%HfPosAH_{i,t}$) on its lagged short interest ($SIR_{i,t-1}$). regression result of

$$\%HfPosAH_{i,t} = \alpha_i + \alpha_t + \delta SIR_{i,t-1} + controls + \epsilon_{i,t}. \tag{A10}$$

In our controls, we include log transformed total net asset value, and year, quarter, and ETF fixed effects. All standard errors are clustered by ETFs and quarters. We run the above regression model on the full sample of our earnings announcements ranging from 1995 to 2016, with the exception of 2007 and 2008.

Panel A:Summary Statistics

	Mean	**Std.**	**5%**	**25%**	**50%**	**75%**	**95%**
%HfPosAH	47.488	13.463	26.863	41.176	48.611	55.402	66.667

Panel B: Regression Results

Dependent variable: *%HfPosAH_(i,t)*	**Ind. ETFs**
SIR_(i,t-1)	-0.2146
	(-0.1654)
log(TNA)	0.3341*
	(1.7009)
Year F.E.	Y
Quarter F.E.	Y
ETF F.E.	Y
Num.Obs.	3942
R-squared	0.2342

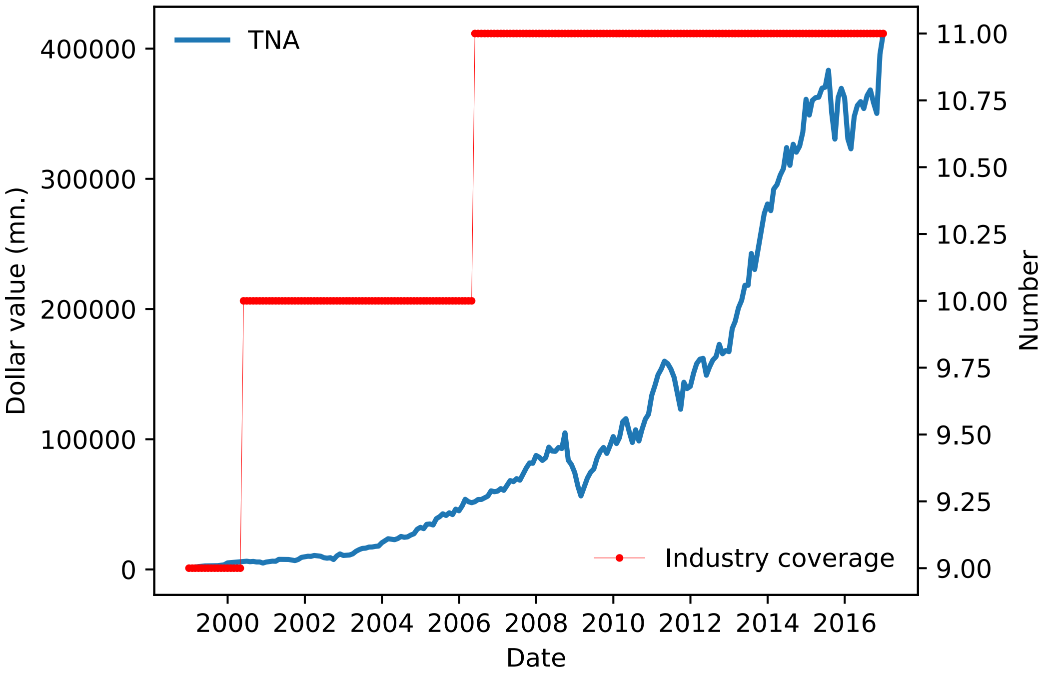

Figure 1: Total net asset value and industry coverage of industry ETFs

Figure 1 shows the time series pattern of total net asset value and industry coverage of industry ETFs from 1998.12.16 (the earliest inception date among our industry ETFs) to 2016.12.31 (the end of our sample period). The blue line is the total net asset value with measurement on the lETF, and the red line is the industry coverage with measurement on the right.

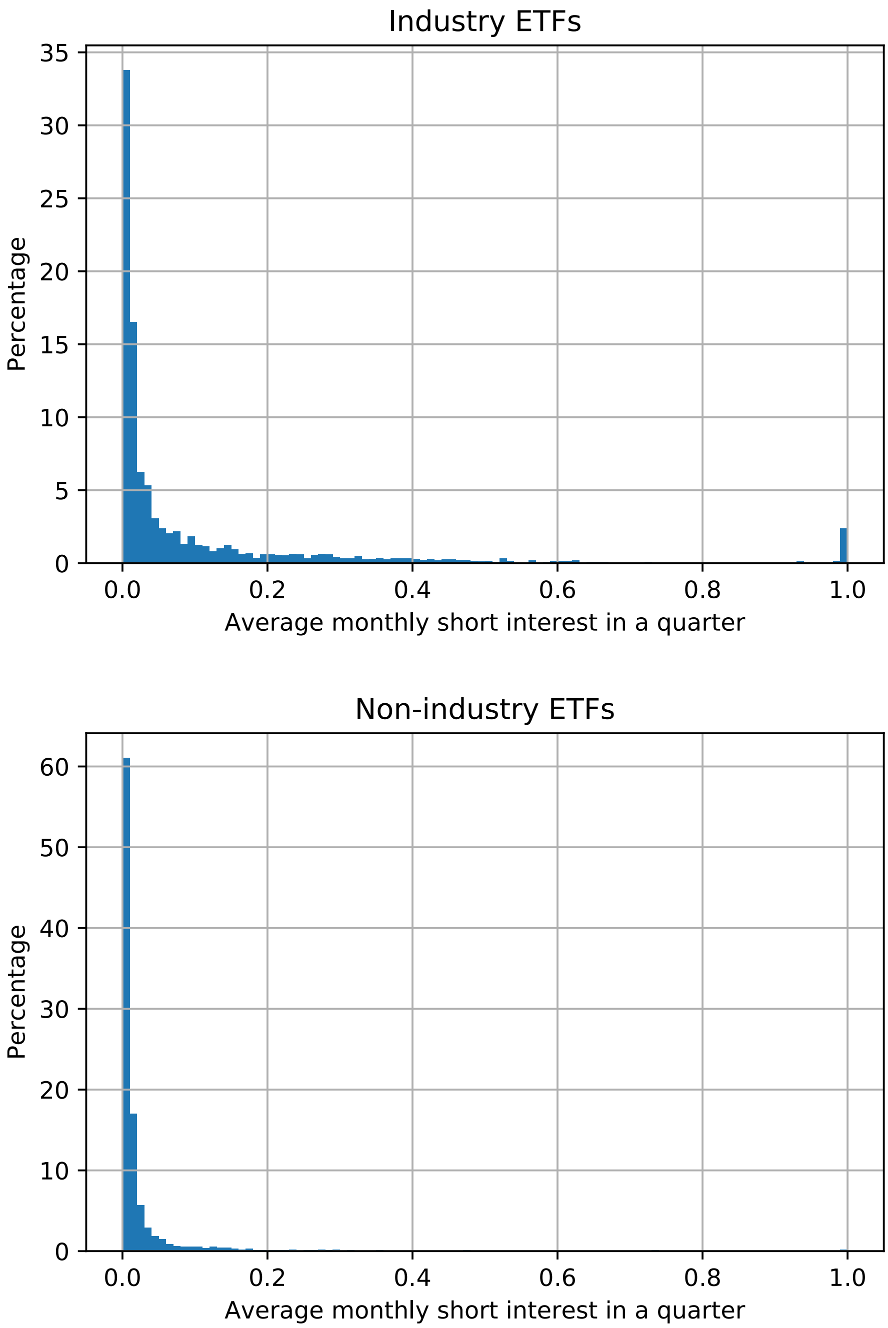

Figure 2: The histogram of the short interest ratio

Figure 2 shows the histogram of the short interest ratio of industry (top panel) and non-industry ETFs (bottom panel), respectively. The short interest ratio is sampled quarterly using the average monthly ration in a quarter.