Market returns and Investor sentiment measured by Internet search volume

Joon Chae

Seoul National University, Seoul, Korea

Hyungjoo Kim

Bank of Korea, Seoul, Korea

Bonha Koo[1]

Seoul National University, Seoul, Korea

[1] Corresponding Author: Bonha Koo, Ph.D Candidate, College of Business Administration, 58-413, Seoul National University, Kwanak Gu, Sinlim-Dong, Seoul, Korea, 151-916; E-mail: koobonha@snu.ac.kr; Tel: +82-2-880-6929.

Market returns and Investor sentiment measured by Internet search volume

Abstract

We propose a new measure of investor sentiment, using weekly Internet search volume data of words in Korea from NAVER DataLab—Financial and Economic Attitudes Revealed by Search (FEARS) index, which reflect households' economic concerns. We find that the FEARS index: (1) relates to a negative contemporary return and reverses after three weeks; (2) coincides with a temporary increase in volatility; (3) simultaneously induces a shift in trading behavior from risky to safe assets, which reverses after three weeks; and (4) is mostly affected by individual investors. Overall, these results are consistent with behavioral finance, which suggests that investor sentiment leads to mispricing that is subsequently corrected.

Keywords Investor sentiment; Internet search volume; Return reversal, FEARS index; NAVER DataLab

JEL Classification: G10

1. Introduction

The efficient market hypothesis states that prices should always be consistent with fundamental values by reflecting all available information. Even after mispricing occurs due to irrational investors, mispricing converges to equilibrium price through rational arbitrageurs. However, various empirical analyses seem to challenge the efficient market hypothesis and actively promote behavioral finance. According to literature on behavioral finance, investors are not fully rational, and their demand for risky assets is affected by sentiments that are not fully justified by fundamental news. Further, arbitrage is risky, and therefore limited. This collectively imply that changes in investor sentiment are not fully countered by arbitrageurs and thus affect security returns (Shleifer and Vishny, 1997). Therefore, it is important to study how investors are affected by psychological factors, and how stock prices reflect them. Baker and Wurgler (2007) emphasize the importance of both quantifying investor sentiment and analyzing its effects.

The main object of this paper is to quantify the Korean market's investor sentiment by analyzing Internet search volume data and to investigate its impact on the stock market. Korea is the one of a few number of countries that investor sentiment can be clearly quantified by Internet search-based methods for the following reasons: first, the country has a unique investor environment, including a high proportion of individual investments in the stock market. According to the Korea Exchange, the proportion of trading volume (made) by individual investors in Korea was approximately 91% in 2015, while household ownership of the U.S. equity market at the end of June 2016 was

37%[2] and 11% of U.K. equity market by value was held by individuals in 2014[3]. Further, almost everyone in Korea uses the Internet, and the nation has an excellent internet infrastructure. Data provided by the Organization for Economic Cooperation and Development (OECD) in 2014 indicates that Korea has had the best Internet access of households in OECD countries for ten consecutive years.[4] The Korea Internet & Security Agency notes that the percentage of Internet users of Korea's total population was 82.10% in 2013. Additionally, Korea is known for providing the world's fastest average Internet speed.[5] Thus, we can assume that Internet search trends effectively reflect the millions of households in Korea.

Korea's leading Internet search engine company, NAVER Corporation,[6] provides a weekly public search volume data via its NAVER DataLab (http://datalab.naver.com/). When a user inputs a search term into the NAVER DataLab, it displays the search volume history scaled by the time-series maximum. Figure 1 plots two examples of NAVER's graphical output of its search volume index (SVI) to note that the Internet search volume reveals household sentiment. The top panel in Figure 1 represents the weekly SVI for "Financial Crisis," the bottom panel represents the weekly SVI for "Depression," and the number of 100 represents the peak search volume over the whole period. The SVI for both "Financial Crisis" and "Depression" began increasing

---

2 Source: Federal Reserve and Goldman Sachs Global Investment Research, as of 2Q 2016.

3 Source: The Office for National Statistics, 2014.

4 According to the OECD's data (https://data.oecd.org/), Korea's Internet access remains consecutive from 2005 to 2014. Internet access is defined as the percentage of households who reported that they had access to the Internet.

5 According to the "state of the Internet report in Q1 2016" by Akamai Technologies (http://www.akamai.com), a leading global company in the content delivery network (CDN) field, South Korea has had the fastest average Internet speed worldwide for nine straight quarters.

6 Following Internet trend (http://www.internettrend.co.kr), the NAVER is the dominant search engine in Korea, with a market share of 84.64% from January 2015 to June 2016, while market share of DAUM is 12.46%, Google is 1.18%, and ZUM is 1.14%.

dramatically in the middle of 2008. Further, this time is consistent with the period in which the Korea Economic Research Institute announced that Korea was in a recession. This roughly suggests that household sentiment could be revealed by SVI.

We verify that SVI correlates with investor sentiment by plotting the monthly log SVI for “Economic Crisis” against the monthly Consumer Composite Sentiment Index (CCSI) in Figure 2. The CCSI is the monthly index measured by survey-based methods, which primarily collects data on households’ economic outlooks from the Bank of Korea (BOK). If the CCSI exceeds (or is less than) 100, this means that many people have optimistic (or pessimistic) views of the future. Therefore, we add a minus sign to the log SVI, as a higher SVI for “Economic Crisis” signals pessimism. Figure 2 confirms that two time series moved similarly, with a correlation coefficient of 67.5%. Accordingly, we can infer that SVI can be a suitable proxy to reflect investors’ sentiments.

As a measure of investor sentiment using the SVI, we employ a Financial and Economic Attitudes Revealed by Search (FEARS) index, following the work of Da *et al.* (2015). We construct the FEARS index using the weekly SVI related to economic words with positive and negative tags, and we choose words that most reveal the households’ economic concerns. A list of search term in FEARS from January 2007 to June 2016 includes not only words with negative sentiment tags, such as “Financial Crisis” and “Recession,” but also words with positive sentiment tags, such as “Gold” and “Stock.”

We use this index to perform several empirical analyses to test whether this reflects investor sentiment. De Long, Shleifer, Summers, and Waldmann (DSSW; 1990) state that if irrational noise traders make trading decisions based on their sentiment, and risk-averse arbitrageurs encounter limits to arbitrage, this will cause more noise trading, greater mispricing, and excess volatility. Thus, we test the relationships between the FEARS index and asset returns, volatility, and fund flows. Furthermore, we examine how the effect of the FEARS index on trading behavior of individual investors differs from that on trading behavior of institutional investors, which is a new approach for search-based investor sentiment paper.

Our empirical results suggest that we can predict future return reversals using the FEARS index. The increases in the FEARS index correspond to contemporaneous decreases in market returns. This negative relationship persists for the first two weeks, and then reversal occurs in the third week. In addition, we observe that reversals of the set of small stocks appear later than the set of large stocks. Moreover, we find a stronger reversal among the portfolio of higher CAPM beta, and this results corresponds with Baker and Wurgler's (2006) argument that reversal patterns are stronger in higher sentiment. Also, we can find that the FEARS sentiment have stronger effect on the set of higher volatility stocks, which is consistent with the view of Baker and Wurgler (2007) that stocks with higher volatility are riskier and consequently more difficult to arbitrage than stocks with lower volatility. Overall, these results are consistent with behavioral finance, which suggests that investor sentiment leads to mispricing that is subsequently corrected.

Further, this demonstrates that excessive volatility induced by sentiment is temporary. We observe that only contemporaneous market volatility, measured by realized volatility, the KOSPI200 volatility index (VKOSPI), total stock return volatility, and VKOSPI future contracts' returns, indicates a significantly positive relationship with FEARS. This supports Black(1986) assertion that uninformed noise trading based on extreme investor sentiment will cause temporal excessive volatility.

Our test results based on the fund flows provide evidence indicating "flight-to-safety," in that investors shift their investment from equity funds to money market funds (MMF) when the FEARS is at its peak, and this reverses in the third week.

Further, when we test trading activity by investor type on FEARS, we find that the KOSPI market demand of individual investors, mostly classified as uninformed irrational investors, moves into the Korea Securities Dealers Automated Quotations (KOSDAQ) market when FEARS is at its peak. As the KOSPI is relatively safer than the KOSDAQ market, this broadly implies a flight-to-safety. Moreover, we can find that only individuals' trading behavior in the KOSPI market has the reversal patterns on the third week, while institutional investors' trading seems less linked to the FEARS index. This implies that our sentiment index mainly related to the uninformed noise traders, not the informed rational investors. These findings are unique to our research and further confirm that our FEARS index is a suitable proxy for investor sentiment. Overall, our empirical results are collectively consistent with investor sentiment theories.

The remainder of the paper is structured as follows: Section 2 describes the literature reviews regarding investor sentiment. Section 3 discusses the data and methodology. Section 4 presents the empirical results of our findings, and Section 5 concludes the study.

2. Literature Review

Traditionally, investor sentiment has been measured in two ways: the market-based and survey-based approaches. Many studies under the market-based approach employ stock-return based proxies. For example, Baker and Wurgler (2006, 2007) use a composite index of sentiment, based on six variations: trading volume, dividend premium, the closed-end fund discount, the number and first-day returns on initial public offerings, and the equity share in new issues. Other studies use micro-trading data. Wang (2001) employs the trading positions of large speculators in the futures markets, whereas Kumar and Lee (2006) use broker and transaction data, respectively. Chung and Kim (2009) in Korea use turnover rate and Jang and An (2012) use the Greed and Fear index (GFI) and VKOSPI. Further, Kim and Byun (2010) and Byun and Kim (2013) employ a composite index following the work of Baker and Wurgler (2006). Consumer surveys (Brown and Cliff, 2004; Qiu and Welch, 2006; Menkhoff and Rebitzky, 2008) and consumer confidence indices (Lemmon and Portniaguina, 2006; Schmeling, 2009) are employed as proxies for sentiment under the survey-based approach. Park (2005) uses the Consumer Confidence Index provided by Statistics Korea.

Recently, various attempts have been made to measure investor sentiment using Internet searches, news, and social network data. Many studies, including those by Da *et al.* (2015), Joseph *et al.* (2011), Preis *et al.* (2013), Beer *et al.* (2013), use Internet search volumes provided by Google Trends to quantify investor sentiment in various ways. Moreover, a few Internet search-based works of literature exist in Korea. Nam *et al.* (2012) use NAVER Finance, a Korean Internet stock-related message board, to show that investment opinions can explain stock returns. Koo and Kim (2015) investigate the relationship between firms' Internet search volume and their stock return and trading volumes. Kim and Koo (2013) study investment strategies using Internet search trends following the work of Preis *et al.* (2013), and demonstrate that no meaningful result outperforms the market's average.

Search behavior revealed by search data, unlike traditional approaches, has the ability to be more objective by using external measurements, which are prone to be less driven by economic phenomenon. Market-based approaches might experience problems, as they use output measures to obtain a model input measure, sentiment, a suggested output measure, and financial mispricing. Therefore, mispricing in financial markets is not driven by sentiment, but by another economic phenomenon (Qiu and Welch, 2006). Further, the search-based sentiment approach has several advantages compared to the survey-based approach, in that it is available at a relatively high frequency. Search-based indexes are often available weekly, while consumer confidence indexes are often available monthly. Moreover, search-based measurements could be more objective; survey answers might be inaccurate as little incentive exists to answer survey questions truthfully or carefully, and especially when questions are sensitive (Singer, 2002).

3. Data and Methodology

3.1 FEARS index

The first step to constructing the FEARS index, which is the main variable, involves establishing a list of search terms that reveal sentiment toward economic conditions. Da *et al.* (2015) use the Harvard IV-4 Dictionary and Lasswell Value Dictionary[7] to select a list of search terms. The authors employ all economic words[8] from these dictionaries with a "positive" or "negative" sentiment tag, which results in 149 words. After adding related terms and deleting those with duplicate or insufficient data in Google Trends, they have a final 118 search terms.

We conducted a lot of consideration in the selection of search terms. We attempted

---

[7] Harvard IV-4 Dictionary and Lasswell Value Dictionary, available from http://www.wjh.harvard.edu/~inquirer/homecat.htm

[8] Economic words have the "Econ@" or "ECON" tag in the Harvard IV-4 Dictionary and Lasswell Value Dictionary.

to obtain a Korean dictionary with a category for economic fields and sentiment tags, as in the above dictionaries, to minimize any divergence in research results due to differences in English and Korean terms, but we were unable to locate one. Previous literature used search terms following papers in the United States; for example, Kim and Koo (2013) use 84 words from a literal translation of Preis *et al.* (2013), which includes a final 98 words. However, Kim and Koo (2013) state that this set of specified words requires a more conscious approach due to the significant difference between environments in the United States and Korea. We consider this, and translate 163 terms with all "positive" or "negative" tags over 502 economic words in the Harvard IV-4 Dictionary and Lasswell Value Dictionary. The results would provide a better fit to Korea by not directly using the final 118 search terms from the work of Da *et al.* (2015). After deleting adjectives and homonyms from the dictionary, we have 140 English words, and after translating these words into Korean, we have 192 Korean words, including all primitive Korean words obtained when English words are inserted into the dictionary. We then input each primitive word into NAVER to include all the words that might be searched in NAVER by households, and include at most ten related search terms for each primitive word. After removing terms with duplicate words or homonyms, we have 365 related search terms; after deleting insufficient data, we have a final 140 search terms.

Next, we collect the SVI data for each of these 140 words from January 2007 to June 2016 from the NAVER DataLab, which provides a weekly search volume[9] for any keyword from January 2007. We define the weekly log change in search term j at time t

[9] As the NAVER DataLab only provides the public weekly SVI, we did not use the daily SVI, as in the work of Da *et al.* (2015).

as:

$$\Delta\mathrm{SVI}_{\mathrm{j,t}} = \ln\left(\mathrm{SVI}_{\mathrm{j,t}}\right) - \ln\left(\mathrm{SVI}_{j,\mathrm{t}-1}\right) \qquad (1)$$

In Figure 3, we plot two examples of $\Delta\mathrm{SVI}_{\mathrm{j,t}}$, “Recession” and “Stock Market” from January 2016 to June 2016. We mitigate any concerns regarding outliers, seasonality, and heteroscedasticity in the sample data by adjusting the $\Delta\mathrm{SVI}_{\mathrm{j,t}}$ as follows. First, all $\Delta\mathrm{SVI}_{\mathrm{j,t}}$ are Winsorized at the 1 and 99 percent levels. We also regress $\Delta\mathrm{SVI}_{\mathrm{j,t}}$ on both month and year dummies and keep the residual. We then standardize $\Delta\mathrm{SVI}_{\mathrm{j,t}}$ by scaling the standard deviation of each time series to make each time series comparable. This creates an adjusted weekly change in search volume, $\Delta\mathrm{ASVI}_{\mathrm{t}}$, for each of our 140 terms. Table 8 illustrates the robustness to unadjusted weekly changes in search volume.

Our final step involves selecting the most useful search terms to represent market returns. We run expanding backward rolling regressions of $\Delta\mathrm{ASVI}$ on market returns every six months, or every June and December, to identify the historical relationship between search term and market return for each of our 140 terms. We use the KOSPI index to measure the market; this data is obtained from FnGuide (www.fnguide.com). Despite the fact that we use both positive and negative sentiment in dictionaries, we find that most search terms have a negative relationship with market return. For example, we find only 3 terms with a t-statistic on $\Delta\mathrm{ASVI}$ greater than 2.5, while 17 terms have a *t*-statistic of less than -2.5 during the sample period. This implies that changes in stock prices are more sensitive to negative information than positive information, as posited by Engle and Ng (1993). Further, this suggests that negative terms are more useful in identifying sentiment (Tetlock, 2007). Therefore, we use only 30 search terms that have the largest negative *t*-statistic on ΔASVI to construct our FEARS index. We define the FEARS for week *t* as:

$$\text{FEARS}_{\text{t}} = \sum_{i=1}^{30} R^{i}\ (\Delta\text{ASVI}_{t}) \ , \tag{2}$$

where $\text{R}^{\text{i}}\,(\Delta\text{ASVI}_{t})$ is the rank *i* of ΔASVI_{t}'s *t*-statistic from the period of January 2007 through the most recent six-month period, and ranks from the most negative (*i* = 1) to the most positive (*i* = 140) *t*-statistic. For example, we run a regression of ΔASVI on contemporaneous market returns during the period from January 1, 2007, to June 30, 2010, for each of our 140 search terms. We then rank the *t*-statistics for this regression, and select the 30 most negative terms for the period from July 1, 2010, to December 31, 2010. The FEARS index for week *t* during this period is the average ΔASVI of these 30 terms on week *t*. We use an expanding rolling window due to the relatively short sample period, and use 30 terms as this is considered the minimum number of observations for diversifying idiosyncratic noise, following the work of Da *et al.* (2015). We perform robustness checks for an alternative cutoff in Table 8. Our FEARS index begins in July 2007, as it requires at least a six-month initial window.

Table 1 reports a list of FEARS terms from our sample period (January 1, 2007, to June 30, 2016) in Panel A, and Panel B notes the United States FEARS terms from the work of Da *et al.* (2015), from January 2004 to December 2011. The terms in Panel A are ordered from the most negative (Financial Crisis) to the least negative *t*-statistic (Illegality). We compare Panels A and B to identify the similarities and differences of the FEARS terms between the markets in Korea and the United States. The Korean FEARS term, as in the United States, contains economic words with negative sentiment, such as "Economic Crisis" and "Depression." Further, we find that "Gold," which is classified as a positive economic word, has a strong, negative relationship with market returns. The existence of “Gold” in our FEARS term list is natural, as Baur and Lucey (2010) state that gold is considered a safe haven in extreme stock market conditions, while a hedge against stocks in average market conditions. Our FEARS term, unlike in the United States, does not contain a word relating to either “donation” or “charity.” Further, it includes such search terms as “Japanese Economy,” “US Dollar,” and “Eurodollar,” indicating a characteristic of South Korea that is inevitably affected by developed countries. Additionally, we find that the economic words that have a strong, positive relationship with market returns include “Stock Commissions” (t-statistic = 4.57) and “Overseas Purchase Tariff” (t-statistic = 2.29).

3.2 Control variables

Qui and Welch (2006) and Da *et al.* (2015) state that we should expect that investor sentiment to be endogenous to macroeconomic conditions. The large spikes in search volume could relate to some macroeconomic events. For instance, Figure 1 show that SVI for both search words, ‘Financial Crisis’ and ‘Depression’, increases dramatically when Korea was in recession. Since some news will affect investor sentiment while some news will not, Da et al. (2015) control for news events to the extent that the rate of returns, the business conditions index and the policy uncertainty index measures. In this way, FEARS index could describe the amount of sentiment generated by an event.

Specifically, Da *et al.* (2015) includes control variables from the Chicago Board Options Exchange’s (CBOE) daily market volatility index (VIX), changes in the

Aruoba-Diebold-Scotti (ADS) business conditions index,[10] and changes in a news-based measure of economic policy uncertainty (EPU) in most specifications.

We use the VKOSPI index as an alternative to VIX, which represents the implied volatility of options on the S&P 100 stock index. Baker and Wurgler (2007) consider it an alternative market sentiment measure. Similar to VIX, VKOSPI is the implied volatility index of options on the KOSPI200, which is called as "fear index." Moreover, we use a coincident composite index (CCI) as a substitute of ADS, which measures current economic conditions by comprising 10 cyclical economic data sets.[11] The change in the CCI reflects innovations driven by macroeconomic conditions, such as ADS. An increase in the CCI index implies progressively above-average conditions, while a decrease in the CCI index implies progressively below-average conditions. As this involves monthly data, we have usage limitations, such as including the same CCI index values for all weeks in each month. We omit the alternative variable of the EPU, as no similar index exists in Korea. Thus, we use control variables of the alternative sentiment measures VKOSPI and CCI and five lags of market returns, following the work of Beer et al. (2013). These control variables are available from FnGuide.

4. Empirical Results

4.1 Stock Returns

Investor sentiment theories suggest that short-term reversals are evidence that market prices may reflect investor sentiment, and stronger price reversals appear in stocks with

---

[10] The ADS index includes macroeconomic variables of weekly initial jobless claims; monthly payroll employment, industrial production, personal income less transfer payments, manufacturing and trade sales; and quarterly real gross domestic product (GDP).

[11] Index of ① labor input ② industrial products ③ the manufacturing operation ratio ④ producer's shipments ⑤ electronics consumption ⑥ wholesale and retail sales returns ⑦ non-durable consumption goods shipments ⑧ cement consumption ⑨ real amounts of exports ⑩ real amounts of imports

more limited arbitrage opportunities. Thus, we examine the relationship between the FEARS and stock returns to find evidence of return reversals, and examine how this relationship varies in limits of arbitrage.

4.1.1 Asset Returns

We test the relationship between the FEARS and contemporaneous, predictive market returns by conducting the following regression:

$$\mathrm{Ret}_{\mathrm{i,t+k}} = \beta_0 + \beta_1 * \mathrm{FEARS}_{\mathrm{t}} + \textstyle\sum_m r_m * Control^{m}_{\mathrm{i,t}} + \mathrm{u}_{\mathrm{i,t+k}} \, , \tag{3}$$

where $\mathrm{return}_{\mathrm{i,t+k}}$ is the asset *i*'s return on week $t + k$. A set of control variables, $Control^{m}_{\mathrm{i,t}}$, includes lagged asset-class returns of up to five lags, the volatility index of the KOSPI200 (VKOSPI), and the changes in the coincident composite index (CCI). A negative coefficient for the FEARS would imply declines in market returns when the search volume increases for FEARS terms, such as "financial crisis" and "stock market."

Table 2 reports the relationships between the FEARS and KOSPI returns by using Equation (3). The KOSPI index has been downloaded from FnGuide, and indicates a significantly negative relationship between the FEARS and KOSPI returns, at k = 0. The coefficient of correlation is -5.66%, and is statistically significant at 1%. This indicates that increases in the FEARS index correspond with contemporaneous decreases in the KOSPI index. This phenomenon seems natural, as we select the FEARS search term based on the historical relationship with KOSPI returns.

We then examine data from the next few weeks seeking evidence of return reversals. We observe a significantly negative relationship between the FEARS and KOSPI in the second week. We consider that the coefficient of FEARS at k = 1 has no significance, and test cumulative returns from week 0 to week 2, or $\mathrm{Ret}(\mathrm{t},\ \mathrm{t}+2)$; this also has a significantly negative coefficient. Column (5) verifies that the FEARS index strongly predicts future return reversals in the third week, which demonstrates a significantly positive relationship between the FEARS index and market returns. The coefficient of correlation is 2.33%, and is statistically significant at 5%. We additionally consider the period until the fifth week, but none of the coefficients of FEARS indicates reversals.

We test regression (3) with different assets in Table 3 to examine that reversals also appear in other asset classes. Panel A illustrates the results from testing the KOSPI200 value-weighted and KOSPI200 equally weighted indexes, as the former has a tendency to reduce the impact of small stocks and the latter has a tendency to reduce the impact of large stocks. Table 2 presents both contemporaneous and cumulative returns until the second week (Columns (1), (2), (4) and (5)) for both indexes, and illustrates the significantly negative coefficient for the FEARS, further predicting a significant return reversal in the third week (Columns (3) and (6)).

Panel B focuses on the large and small stock indexes obtained from FnGuide. The large stock index is a value-weighted index, which includes the top 100 market capitalization stocks among the MKF 500[12] stocks; the small stock index is the bottom 300 market capitalization stocks among the MKF 500 stocks. Contemporaneous and cumulative returns until the second week in both indexes reveal the FEARS' negative coefficient. Interestingly, the small stock index does not predicts a return reversal in the third week, while the large stock index predicts a return reversal in the third week. Regarding small stock returns, an unreported coefficient for the FEARS in the fourth

[12] The MKF500 is an index that includes the top 500 market capitalization of stocks listed on the KOSPI and KOSDAQ markets.

week is 1.95%, and is statistically significant at 10%. The one-week delay in small stock reversals can be explained by information asymmetry, as small stocks' information is prone to reflect more slowly in the stock prices than large stocks. Alternatively, large stocks' information is reflected more quickly in their prices, as information is provided more frequently to the market by analysts and investors due to its impact on the market. Lo and MacKinlay (1990) argue that the returns on large stocks lead those on smaller stocks, while an opposite phenomenon does not occur. Further, Conrad *et al.* (1991) posit that large stock returns' volatility is used to predict that of small stocks, while an opposite trend is not observed. These results prove that a difference exists in the speed of this information being reflected between large and small stocks.

Panel C demonstrates test results using the KOSDAQ and the treasury-bond index. No evidence exists of reversals on the KOSDAQ, although we find a significantly negative coefficient for contemporaneous and cumulative returns until the second week. This seems to be caused by relatively insufficient information, such as in the small stock index. We use a total return index of MKF treasury bonds, obtained from FnGuide, to measure the treasury-bond index. The treasury-bond index, contrary to other asset classes, exhibits a significantly positive relationship between the FEARS index and cumulative market returns until the second week (Column (5)). This is broadly consistent with the flight-to-safety concept, in that investors prefer treasuries as a safe haven during times of increased uncertainty.

Results in Tables 2 and 3 demonstrate that the FEARS index strongly corresponds with contemporaneous and cumulative returns until the second week, then predicts future return reversals in the third week (or the fourth week in case of small stocks).

Regarding the work of Da *et al.* (2015), the FEARS for the United States' data negatively corresponds with contemporaneous market returns, and predicts reversals in the forthcoming two days. Although we cannot compare our weekly data results to the United States' results on a daily basis, it seems that our data's reversals occur later than in the United States. This might be caused by the limits of arbitrages due to institutional differences between the two countries, such as the prohibition of short sales and the restrictions of price ranges on the Korean stock market.

Further, these relatively later reversals might be explained by a liquidity shock. Yang (2010) states that market liquidity (KOSPI) decreases much more in a declining market, related to its increases in a rising market. This also indicates that a 1% decrease in market returns significantly increases 0.026% of bid-ask spreads for the cumulative two weeks. Additionally, according to Baker and Stein (2004), a liquidity shock can be a sentiment indicator in a world with short-sales constraints. The authors explain the lower subsequent returns after high liquidity by irrational investors who boost this high by underreacting to information. Therefore, our results are consistent with investor sentiment theories, even if this phenomenon is caused by liquidity shocks.

4.1.2 Limits of arbitrage

DDSW (1990) argue that investors are of two types: rational arbitrageurs who are sentiment-free and irrational traders prone to exogenous sentiment. And mispricing arises out of the combination of two factors: irrational traders make trading decisions based on their sentiment and risk-averse arbitrageurs encounter limits to arbitrage. Limits to arbitrage is one of the most important channels that can exacerbate the effect of investor sentiments (Shleifer and Vishny, 1997).

Motivated by limits to arbitrage, we estimate the effect of sentiment on portfolio return spreads constructed by sorting on stock characteristics related to limits to arbitrage. One motivation for the tests is to explore the limits of arbitrage view of Baker and Wurgler (2007) which argue that stocks with higher volatility are riskier and consequently more difficult to arbitrage than stocks with lower volatility. So, we use the volatility of stock returns as a measure of the difficulty of arbitrage. We examine the FEARS effect on the return spreads between high-volatility and low-volatility stock portfolios. We estimate the volatility with a return of common stocks included in the KOSPI within a year and then sort our sample into decile portfolios based on the volatility.

We also consider the test on return spreads between high-beta stock portfolios and low-beta stock portfolios, motivated by Baker *et al.* (2011) which argue that high-beta portfolios are easy subjects of speculation of irrational investors and high beta stocks may not be attractive to arbitrageurs who have institutional constraints such as benchmarking. By this, we could assume that investor sentiment have a larger impact among high-beta stocks than among low-beta stocks. In this hypothesis, the return spreads should be negatively related to contemporaneous increase in FEARS index, while future return spreads should be positively related to FEARS index. Therefore, we test the relationships between the FEARS and return spreads from beta-sorted portfolios by using Equation (3) in Panel A of Table 4. We estimate the CAPM beta by the following equation:

$$R_{\mathrm{it}} - R_{ft} = \alpha_i + \beta_i \left(R_{mt} - R_{ft}\right) + \varepsilon_{it}, \tag{4}$$

where R_{it} is the excess return to stock i at time t, R_{mt} is the (KOSPI) market return at time t, R_{ft} is the risk-free rate of return(364 day's monetary stabilization bond) at time t, and $\tilde{\mathcal{E}}_{it}$ is the residual. We use the daily return data of common stocks included in the KOSPI within a year. We avoid liquidity concerns by restricting our sample to stocks included in the KOSPI, and employ the weekly average beta for each stock and form decile portfolios. We then compute the high-minus-low return spread with these beta-sorted decile portfolios.

The results in table 4 are consistent with our investor sentiment hypothesis. We can observe that the FEARS has a more negative contemporaneous relationship with the high beta stock (Columns (1)). The coefficient on the FEARS at k = 0 is -3.86%, and is statistically significant 5%. Moreover, as evident in columns (3), an increase in FEARS on this week (i.e., k =0) predicts a return reversal on the third week (i.e., k =3). The coefficient on the FEARS at k = 3 is 4.73%, and is statistically significant 5%. Similar results are obtained using the high-minus-low volatility portfolio return spreads. We can find that the FEARS sentiment have stronger effect on high-volatility stocks than low-volatility stocks on week t (Columns (4)), while the impact reverses on the third week (Columns (6)).

Next, we test the relationship between the FEARS and the downside risk in Table 4, following the work of Da *et al.* (2015). Ang et al. (2006) state that if an asset tends to move downward in a declining market more than it moves upward in a rising market, it is an unattractive to arbitrageurs because it tends to have very low payoffs. Since downside risk limits arbitrageurs from correcting mispricing, we can assume that stocks with high downside risk underperform the stocks with lower downside risk when downside risk is large and investor sentiment is high.

We use two measures of downside risk, downside beta and downside volatility, following the work of Ang *et al.* (2006). First, we estimate the downside beta (β_i^-) for individual stocks by using following equation:

$$\beta_i^- = \frac{cov(r_i, r_m | r_m < \mu_m)}{var(r_m | r_m < \mu_m)} \ , \tag{5}$$

where r_i is security i's excess return, r_m is the market's (KOSPI) excess return, and μ_m is the average market excess return. We use the weekly average downside beta for each stock to form downside beta-sorted decile portfolios, and use these to compute the high-minus-low downside beta return spread.

Also, we estimate the downside volatility (σ_i^-) for individual stocks by using following equation:

$$\sigma_i^- = \sqrt{var(r_i | r_m < \mu_m)} \ , \tag{6}$$

We use the weekly average downside volatility for each stock to form downside volatility-sorted decile portfolios, and use these to compute the high-minus-low downside volatility spread.

In Table 4, columns (7) to (9) relate FEARS and high minus low return spreads between the downside beta stock portfolios. Also the high minus low return spreads between the returns of downside volatility stock portfolios are in columns (10) to (12) of Table 4. The test results associated with the downside beta-sorted decile portfolios and downside volatility-sorted decile portfolios, coefficient on the FEARS at k = 3 is 3.23% (significant at the 10% level) and 3.09% (significant at the 5% level), respectively. It indicates that the FEARS have stronger effect on high-downside risk stocks than low-downside risk stocks on week t, while the impact reverses in the third week.

Overall, we find stronger evidence of temporary deviation from fundamentals among the set of stocks with a higher beta, a higher volatility, a higher downside beta or a higher downside volatility in Table 4. Our test results motivated by limits to arbitrage confirms investor sentiment theories of Baker and Wurgler (2007) and Baker et al. (2011)

4.2 Volatility

Black (1986) states that uninformed noise trading based on extreme changes in investor sentiment will temporarily cause more noise trading, mispricing, and excessive volatility. We test the relationships between the FEARS and market return volatility in Table 5 to find evidence of a temporal increase in volatility.

This paper employs four stock market volatility measures. The first measurement is realized volatility, calculated following the work of Gospodinov *et al.* (2006):

$$\mathrm{RV}_{t,\tau} = \sqrt{250} * \sqrt{\left(\frac{1}{\tau}\right) * \sum_{i=t+1}^{\tau} r_i^{\ 2}} \ , \tag{7}$$

where r_i is the log return of daily close prices from the KOSPI index and τ is 22. The second measurement is the VKOSPI index, a proxy for implied volatility. Further, we use total volatility, which is a standard deviation of common stock daily returns listed on the KOSPI within a year. Table 5 displays the results from following regression:

$$\mathrm{Vol}_{i,t+k} = \beta_0 + \beta_1 * \mathrm{FEARS}_t + \sum_m r_m * \mathrm{Control}_{i,t}^m + u_{i,t+k} \ , \tag{8}$$

where $\mathrm{Vol}_{i,t+k}$ is the weekly average volatility estimations for each measurement, and the control variable ($Control_{i,t}^m$) denotes the changes in the coincident composite index (CCI).

The last market volatility measure is the VKOSPI futures contract returns, or the tradable asset returns based on volatility. We can avoid potential econometric issues associated with the above volatility measures, and have a clear interpretation of such, by using VKOSPI future returns. We consider this by conducting a regression using Equation (3), although it has had relatively small observations since VKOSPI futures were first imposed in Korea in October 2014. We computed the weekly return as the change in log prices, using the contract closest to maturity. We use the second closest-to-maturity contract data if the contract closest to maturity has less than four trading dates to avoid measurement errors.

Table 5 confirms a significantly positive relationship between the FEARS and only the contemporaneous market volatility. This result supports the works of Black (1986) and DSSW (1990), in that excessive volatility induced by sentiment is temporary.

4.3 Fund flows

A temporary price deviation from fundamental value is caused by irrational investors' noise trading. To examine the effect of noise traders on the sentiment, we examine the relationships between the FEARS and fund flows for equity, bonds, and MMF in Table 6. We exclude any hybrid funds for a clear interpretation. Fund flows are calculated by each group's change in log net asset value (NAV). The NAV data is available from FnGuide.

$$\mathrm{flow}_{i,,t} = \ln\left(\mathrm{NAV}_{j,t}\right) - \ln\left(\mathrm{NAV}_{j,t-1}\right) \tag{9}$$

We run regressions for fund flows for up to four weeks following the regression:

$$\mathrm{flow}_{\mathrm{i,t+k}} = \beta_0 + \beta_1 * \mathrm{FEARS}_{\mathrm{t}} + \sum_{\mathrm{m}} \mathrm{r}_{\mathrm{m}} * Control_{\mathrm{i,t}}^{m} + \mathrm{u}_{\mathrm{i,t+k}} \,, \tag{10}$$

where fund class *i* includes equity, bond, and MMFs. Control variables ($Control_{\mathrm{i,t}}^{m}$) include the lagged market returns for up to five lags, the KOSPI200 volatility index (VKOSPI), and the changes in coincident composite index (CCI).

We discover evidence of noise trading in Table 6. Panel A illustrates a significantly negative relationship between the FEARS and equity fund flows at k = 0 and k = 2, and this reverses at k = 3. This indicates that investors withdraw their money from equity funds on the week with high negative sentiment, then reinvest in the third week. Conversely, MMFs in Panel C inflow on a contemporaneous week when the FEARS is at its peak and reverses in the third week significant at the 5% level. The MMFs are usually classified as safe assets, as they typically invest in short-term debt with high credit quality, such as treasuries. This implies a flight-to-safety, which indicate that increasing economic uncertainty could change investors' preferences towards a particular investment class, which leads the exit from risky to safe assets. Investors prefer MMFs to equity funds when negative sentiment is high. Moreover, this reverses in the third week; this fund flow reversal after the spike in the FEARS is only demonstrated in our paper.

Regarding bond flows, despite its significantly negative relationship with the FEARS at k = 2 and k = 4, no significant relationship exists between the FEARS and bond flows on a contemporaneous week. This result differs from that of Da *et al.* (2015), which indicates that investors shift their money from equity to bond funds in the contemporaneous week. This phenomenon can be explained by the following reasons. First, Da *et al.* (2015) use only medium-term treasury bonds as their bond funds sample, but we use all pure bond funds in our sample. Further, the portion of individual investors in mutual funds differs; individual investors hold approximately 50% of total mutual funds, while those in the United States approximate 90%. The Korea Financial Investment Association notes that on June 30, 2016, individual investors' portion of bond funds approximates 30%, while the portion in equity funds is approximately 80% in Korea. This suggests that flows of bond funds are not suitable to measure the noise trading of investor sentiment, as individual investors generally induce the latter (Lee *et al.*, 1991).

Overall, we confirm that the FEARS can predict fund flows. When negative sentiment is high, a significant inflow occurs to MMFs, and outflow to equity funds, which reverses in the third week.

4.4 Trading behavior

In addition to the work of Da *et al.* (2015), we examine the trading behaviors of three investor groups (individuals, institutions, and foreigners) to confirm that the FEARS is a suitable proxy for investor sentiment.

Regarding Lewellen, Schlarbaum, and Lease (1974), individual investors typically fail to diversify, holding instead a single stock or a small number of stocks. When investors do diversify, they invest their money to stock-picking mutual funds that charge them high fees while failing to beat the market (Jensen 1968). Black (1986) believes that such investors, with no access to inside information, irrationally act on noise as if it were information that would give them an edge. Black (1986) calls such investors 'noise traders' and argues that noise traders' trading based on extreme changes in investor sentiment will temporarily cause more noise trading. In the sense of Black

(1986), we define individual investors as noise traders, while we define institutional and foreigner investors as informed rational investors who large-scale invest in a group with long-term perspectives.

If noise traders are more sensitive to the FEARS index than other investors, this would imply that the investor sentiment influences the behavior of noise traders. Therefore, we analyzed the relationship between the FEARS and the trading behaviors of each investor type: individuals, institutions, and foreigners. We use the net purchase ratio (NPR) for each investor type to measure trading activity, and calculate the NPR as the ratio of net purchase amount to total transaction amounts, following the work of Lakonishok *et al.* (1992):

$$\mathrm{NPR_{i,t}} = (\mathrm{Buy_{i,t}} - \mathrm{Sell_{i,t}}) / (\mathrm{Buy_{i,t}} + \mathrm{Sell_{i,t}}) \quad (11)$$

If the NPR is positive (or negative), then investors' excess demand *i* at time *t* is positive (or negative). Furthermore, motivated by the results from the section 4.3, which shows the evidence of a flight-to-safety, we divide our market into two classes, KOSPI and KOSDAQ. All the data necessary for calculation is determined by FnGuide. Table 7 relates the FEARS to the NPR of three investor types in the KOSPI and KOSDAQ markets. The NPR is calculated for each group by the following regression:

$$NPR_{i,t+k} = \beta_0 + \beta_1 * FEARS_t + \sum_m r_m * Control_{i,t}^m + u_{i,t+k} \quad (12)$$

The independent variable is the FEARS and a set of control variables; $Control_{\mathrm{i,t}}^m$ includes lagged KOSPI returns, for up to five lags; the KOSPI200 volatility index (VKOSPI); and changes in the coincident composite index (CCI).

Table 7 reveals that the FEARS can predict noise traders' future trading behaviors. The Panel A show that only individuals' trading behavior in the KOSPI market has the reversal patterns on the third week. As shown in Panel A, the coefficient for the FEARS is positive and statistically significant for week $t + 0$ and it reverses at $t + 3$. This is opposed to the results from Panel B and Panel C, which show that institutional and foreigner investors do not predict future reversals. To be specific, institutional investors are hardly affected by our sentiment measure. This implies that investor sentiment influences the behavior of noise traders.

Furthermore, we can find that there is a negative and contemporaneous relationship between the FEARS and trading behavior of individuals in the KOSDAQ market in the column(5) of Panel A, which is the opposite results from the trading behavior of individuals in the KOSPI market in the column(1). As the KOSPI is relatively safer than the KOSDAQ market, this broadly implies a flight-to-safety, in which individual investors' KOSPI market demand moves into the KOSDAQ market. These findings are unique to our research and further confirm that our FEARS index is a suitable proxy for investor sentiment.

4.4 Robustness checks

We verify the FEARS' robustness in Table 8, as its construction required several choices. First, we use an adjusted weekly change in search volume ($\Delta SVI_{j,t}$) for each of our 140 terms. A potential concern about applying winsorization and deseasonalization in constructing the FEARS index is that this could create a forward-looking bias. We address this concern in the first columns of Panel A, which report the results from using the FEARS constructed without winsorization and seasonalization. Further, we perform an expanding backward rolling regression for the $\Delta SVI_{j,t}$ on market returns every six months. Rebuilding every six months could impact the results; thus, we rebuild the

FEARS every three months in Panel A.Both results are slightly better than those in Table 2.

Another consideration of the FEARS is whether it is influenced by excessive market returns, although our control variables include lagged market returns for up to five lags. Therefore, we divide the full sample's KOSPI returns into ten groups, and we add dummy variables for the group that demonstrates the highest market return. Additionally, we add holiday dummy variables, as a possibility exists that public holidays affect the increase in search volume. If more than three holidays exist in a week, we add holiday dummies. Both the coefficients and significance are similar to those in Table 2.

Panel C considers robustness by using the top 25 or 35 terms when selecting the FEARS term, while we use the 30 search terms whose $\Delta SVI_{j,t}$ negatively correlates with market returns. Both results are similar to those in Table 2. Generally, all robustness checks exhibit results similar to the original table.

5. Conclusions

This paper proposes a new measure of Korean investor sentiment based on the internet search volume from Naver DataLab and investigate its impact on the stock market. Regarding Internet search volume, we generate a Financial and Economic Attitudes Revealed by Search (FEARS) index as our sentiment measure, following the work of Da et al. (2015).

Our analysis yields several important results. First, this paper show that investor sentiment contributes to predict short-term market returns. We find that an increase in our sentiment index correlates with contemporaneous decreases in stock market returns, and reversal occurs in the third week. In additional, we observe that the reversal of small stocks appears one week later than large stocks, which can explained by information asymmetry. Second, stock market volatility temporarily increases with the FEARS, which supports Black (1986) that excessive volatility induced by sentiment is temporary. Further, we can find that FEARS can predict fund flows. Fund flows shift from equity to MMFs when the FEARS is high, and this reverses after three weeks. Lastly, our results imply that our sentiment indicator influences the trading behavior of noise traders. We can predict that individuals' net purchase ratio in the KOSPI market has the reversal on the third week. Moreover, the clear reversal patterns are only shown in individual investors, while institutional investors, usually classified as informed investor, are hardly affected by our sentiment index. These findings are unique to our research and further support the assertion that the FEARS index is a reliable sentiment measure in Korea.

The FEARS index can directly reflect millions of households' Internet search behavior, considering that almost everyone in Korea uses the Internet. As research using Internet search terms has not been actively conducted in Korea, this paper can suggest a new perspective and motivation for related research. Specifically, compelling future research could determine whether the FEARS can explain multiple market anomalies.

References

Ang, A., J. Chen, and Y. Xing, 2006, Downside risk, Review of Financial Studies, 19, pp. 1191–239.

Baker, M., and J. Stein, 2004, Market liquidity as a sentiment indicator. Journal of Financial Markets, 7, pp. 271–99.

Baker, M., and J. Wurgler, 2006, Investor sentiment and the cross-section of stock returns. Journal of Finance, 61, pp. 1645–80.

Baker, M., and J. Wurgler, 2007, Investor sentiment in the stock market. Journal of Economic Perspectives, 21, pp. 129–51.

Baker, M., B. Bradley, and J.Wurgler, 2011, Benchmarks as limits to arbitrage: Understanding the low volatility anomaly. Financial Analysts Journal, 67, pp. 40–54.

Baur, D. G., and B. M. Lucey., 2010, Is gold a hedge or a safe haven? An analysis of stocks, bonds, and gold. Financial Review, 45, pp. 217–229.

Beer, F., F. Hervé, and M. Zouaoui, 2013, Is big brother watching us? Google, investor sentiment and the stock market, Economics Bulletin, 33, 1, pp. 454-466

Black, F, 1986, Noise. Journal of Finance, 41, pp. 529–543.

Brown, G. and M. Cliff, 2004, Investor sentiment and the near-term stock market, Journal of Empirical Finance, 11, pp. 1-27.

Byun, J. and K, Kim. 2013, Application of Investor Sentiment Index in Financial Studies , The Korean Journal of Financial Management, 30, 4, pp. 225-248.

Chung, C.-H., and S.-K. Kim, 2009, The Linkages between Stock Returns and Market Liquidity As a Measure of Investor Sentiment, The Korean Journal of Financial Engineering, 8, 4, pp. 65-90.

Conrad, J., Gultekin N., and G. Kaul, 1991, Asymmetric Predictability of Conditional Variances, ReviewofFinancial Studies, 4, 4, pp. 597-622.

Da, Z., J. Engelberg, and P. Gao, 2015, The Sum of All FEARS Investor Sentiment and Asset

Prices. Review of Financial Studies, 28, pp. 1-32.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J.Waldmann, 1990, Noise trader risk in financial markets. Journal of Political Economy, 98, pp. 703–38.

Engle, R.F. and V.K. Ng, 1993, Measuring and Testing the Impact of News on Volatility, Journal of Finance, 48, pp. 1749-1778.

Gospodinov, N., Gavala, A., and Jiang, D., 2006, Forecasting volatility. Journal of Forecasting, 25, 381–400

Jang, S. W., and S.C. An, 2012, Investor Sentiment and the Mean-Variance Relation, Journal of Business Research, 27, 3, pp. 63-85.

Joseph, K., M. B. Babajide, and Z. Zhang, 2011, Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. International Journal of Forecasting, 27, pp. 1116-1127

Kim, K. and J. Byun, 2010, Effect of investor sentiment on market response to stock split announcement, Asia-Pacific Journal of Financial Studies, 39, pp. 687-719.

Kim, M. and P. Koo. 2013, A Study on Big Data Based Investment Strategy Using Internet Search Trends, Journal of the Korean Operations Research and Management Science Society, 38, 4, pp. 56-63

Koo, P. and M. Kim, 2015, A Study on the Relationship between Internet Search Trends and Company's Stock Price and Trading Volume, The Journal of Society for e-Business Studies, 20, 2, pp. 1-14

Kumar, A. and C. Lee, 2006, Retail investor sentiment and return comovement, Journal of Finance, 61, pp. 2451-2486.

Lakonishok, J., A. Shleifer, and R. W. Vishny, 1992, The Impact of Institutional Trading on Stock Prices, Journal of Financial Economics, 32, pp. 23-43.

Lee, C., A. Shleifer, and R. H. Thaler. 1991, Investor Sentiment and the Closed-End Fund Puzzle. Journal of Finance 46, 1, pp. 75–109.

Lemmon, M. and E. Portniagunia, 2006, Consumer confidence and asset prices: Some empirical evidence, Review of Financial Studies, 19, 4, pp. 1499-1529.

Lewellen, Wilbur G.; Schlarbaum, Gary E.; and Lease, Ronald C. The Individual Investor: Attributes and Attitudes. J. Finance 29, pp. 413- 433.

Lo, A., and A. C. MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, Review of Financial Studies, 3, 2, pp. 175-206.

Menkhoff, L. and R. Robitzky, 2008, Investor sentiment in the US dollar: longer-term, nonlinear orientation PPP, Journal of Empirical Finance, 15, 3, pp. 455-467.

Nam, D., J. Park, M. Kim, H. Jo and S. Kim, 2012, A Study about correlation between collective intelligence on the Internet stock message board and stock market, Journal of internet electronic commerce research, 12, 2, pp. 149-164

Park, J., 2005, Consumer Confidence, Investor Sentiment and Stock Returns, Korean Journal of Money & Finance, 10, 2, pp. 199-224.

Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk, Journal of Accounting and Economics, 42, pp. 35–52

Preis, T., H.S. Moat, and H.E. Stanley, 2013, Quantifying trading behavior in financial markets using Google Trends, Scientific Report, Rep. 3, 1684.

Qiu, L., and I.Welch., 2006. Investor sentiment measures. Working Paper, Brown University.

Schmeling, M., 2009, Investor sentiment and stock returns: Some international evidence, Journal of Empirical Finance, 16, 3, pp. 394-408.

Shleifer, A., and R.W. Vishny., 1997, The limits of arbitrage, Journal of Finance 52, 1, pp. 35–55.

Singer, E, 2002, The use of incentives to reduce nonresponse in household surveys. Working Paper, University of Michigan.

Tetlock, P. C, 2007, Giving content to investor sentiment: The role of media in the stock market. Journal of Finance, 62, 3, pp. 1139–1168.

Wang, C., 2001 Investor sentiment and return predictability in agricultural futures markets, Journal of Futures Markets 21, pp. 929–952.

Yang, C.W., 2010, A Study on Determinants of Market Liquidity in the Korean Stock Market, Korean Journal of Financial Studies, 39, 1, pp. 103-132.

Figure 1. Illustrations of NAVER Search Volume Index (SVI)

The figures show the graphical output of weekly search volume index (SVI) from the NAVER DataLab (http://datalab.naver.com/). The top panel represents the SVI for the term "Financial Crisis," and the bottom panel represents the SVI for the term "Depression." The plotted SVI is scaled by the maximum over the full sample.

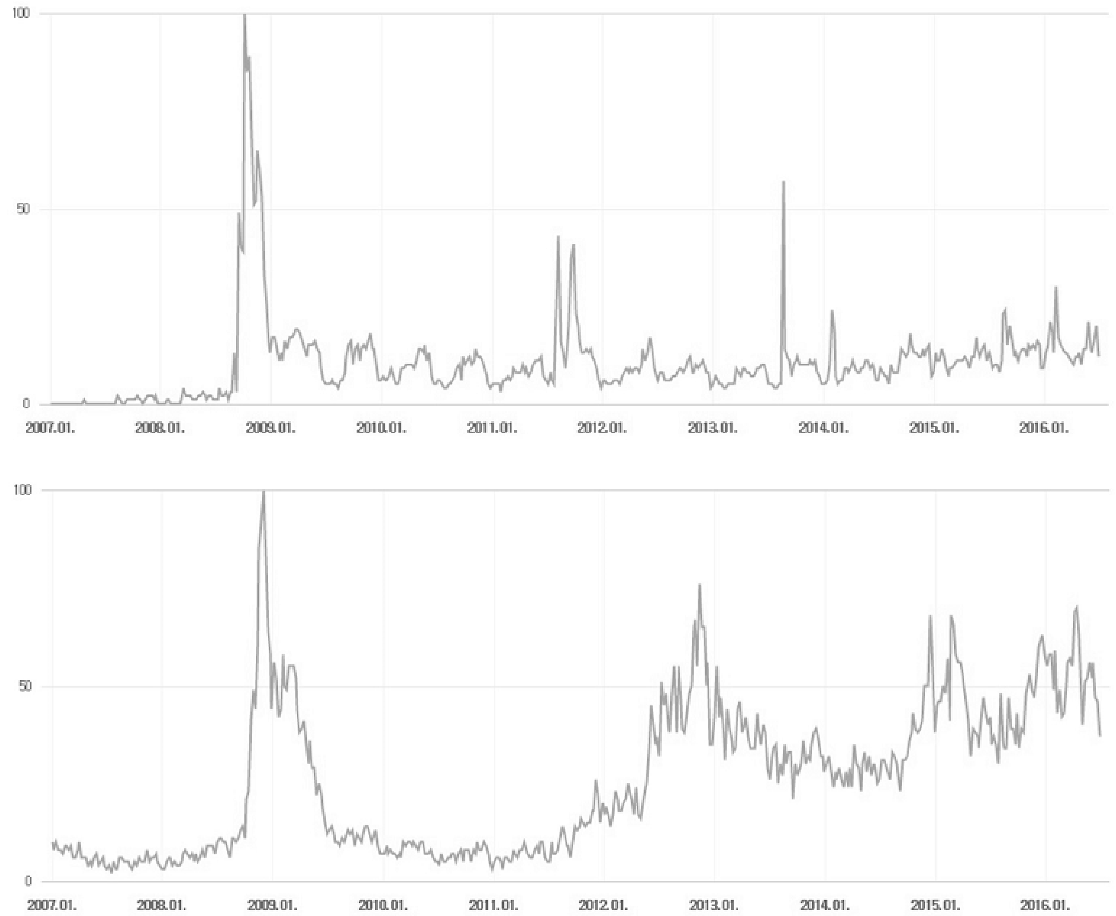


Figure 2. SVI for "Economic Crisis" and Consumer Sentiment

This figure compares the monthly adjusted log SVI for "Economic Crisis" to the monthly Consumer Composite Sentiment Index (CCSI) from July 2008 to June 2016. The correlation between the two series is 0.675.

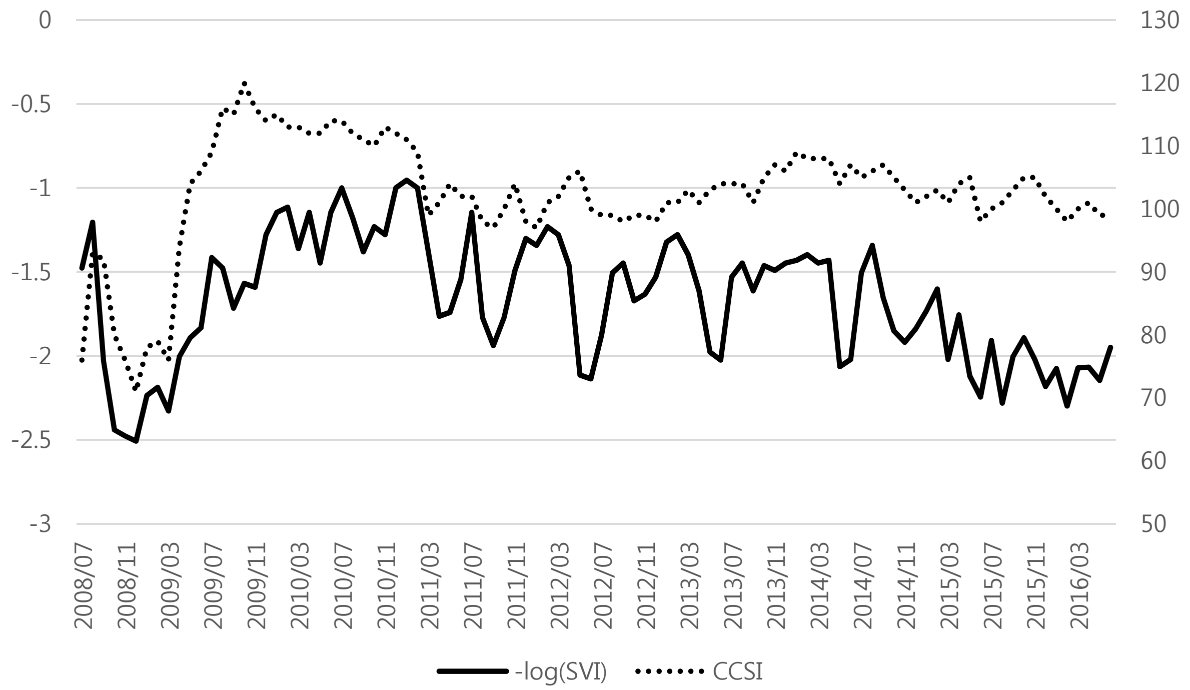
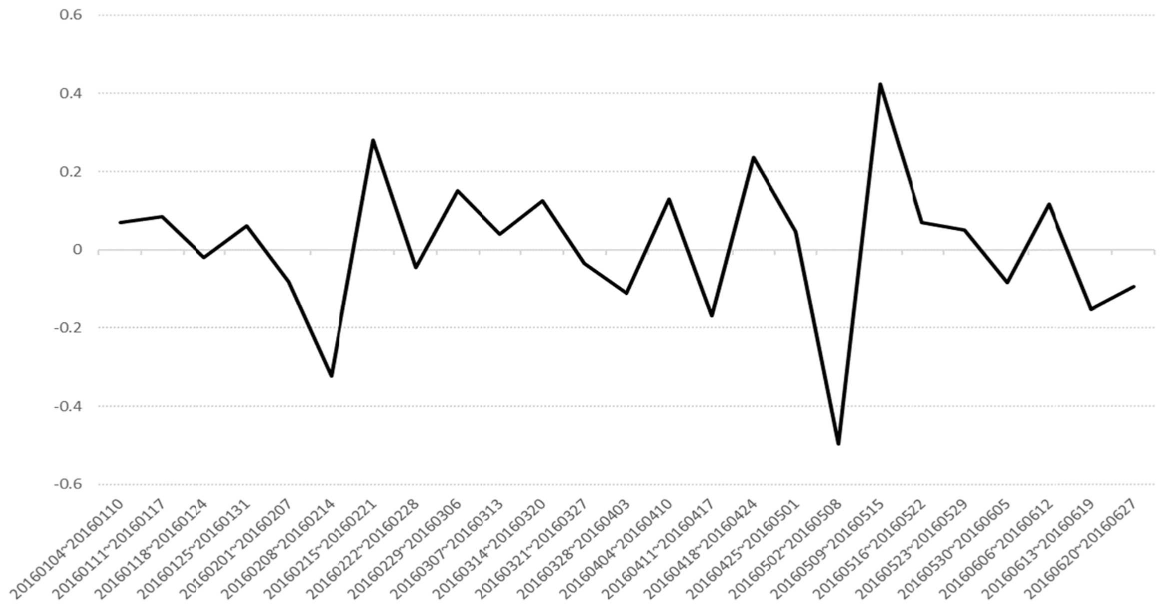


Figure 3. Changes in SVI

The graphs display the weekly log changes for two SVI in Korea over the period of January to June 2016. The top panel presents the weekly change in SVI for "Recession," and the bottom panel presents the weekly change in SVI for "Stock Market."



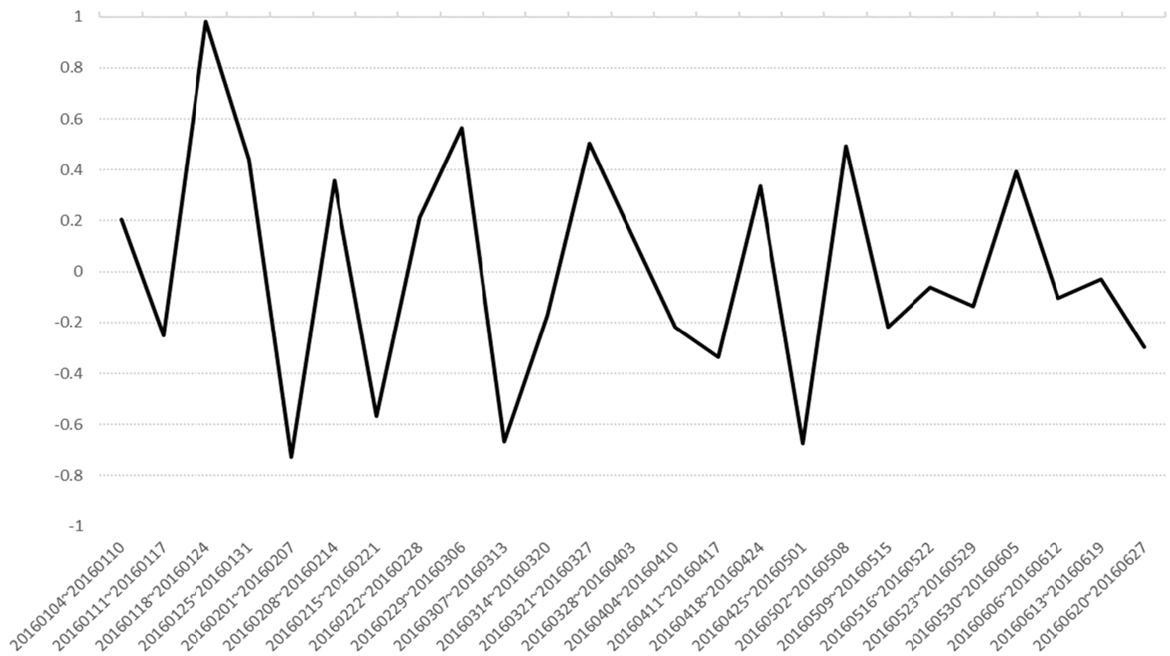


Table 1. The FEARS index search terms

This table displays the top 30 search terms derived from words of economic sentiment with the largest negative market correlation. Panel A indicates the FEARS index search terms from our full sample (January 1, 2007, to June 30, 2016), and Panel B displays the 30 United States FEARS index search terms from January 2004 to December 2011, as noted in the work of Da *et al.* (2015).

Panel A. FEARS terms from the full sample in Korea.

No.	Search Term	T-statistics
1	FINANCIAL CRISIS	-7.25
2	ECONOMIC CRISIS	-6.57
3	STOCK MARKET	-5.23
4	RECESSION	-5.15
5	SECURED LOAN	-4.52
6	DEPRESSION	-3.84
7	STOCK	-3.76
8	PURE GOLD	-3.7
9	SECURITIES MARKET	-3.7
10	GOLD	-3.63
11	U.S. ECONOMY	-3.51
12	INFLATION	-3.31
13	SUBPRIME MORTGAGE	-3.2
14	BOND	-2.51
15	STOCK FUND	-2.44
16	HOMELESS	-2.43
17	TRIGGER	-2.22
18	BUSINESS DEPRESSION	-2.2
19	CRISIS	-2.18
20	JAPANESE ECONOMY	-2.06
21	ECONOMIC DEPRESSION	-2.03
22	U.S. DOLLAR	-2.02
23	EURODOLLAR	-1.75
24	INFLA (SHORTHAND OF INFLATION)	-1.66
25	THE RICH	-1.63
26	BANK DEPOSIT INTEREST RATE	-1.56
27	TIME DEPOSIT INTEREST RATE	-1.48
28	BANK WITH HIGH DEPOSIT INTEREST RATE	-1.36
29	CORRUPTION	-1.36
30	ILLEGALITY	-1.22

Panel B. FEARS terms from January 2004 to December 2011 in the United States.

No.	Search Term	T-statistics
1	GOLD PRICES	-6.04
2	RECESSION	-5.60
3	GOLD PRICE	-4.81
4	DEPRESSION	-4.56
5	GREAT DEPRESSION	-4.15
6	GOLD	-3.98
7	ECONOMY	-3.52
8	PRICE OF GOLD	-3.23
9	THE DEPRESSION	-3.20
10	CRISIS	-2.93
11	FRUGAL	-2.87
12	GDP	-2.85
13	CHARITY	-2.63
14	BANKRUPTCY	-2.50
15	UNEMPLOYMENT	-2.46
16	INFLATION RATE	-2.32
17	BANKRUPT	-2.28
18	THE GREAT DEPRESSION	-2.17
19	CAR DONATE	-2.11
20	CAPITALIZATION	-2.10
21	EXPENSE	-1.97
22	DONATION	-1.89
23	SAVINGS	-1.82
24	SOCIAL SECURITY CARD	-1.71
25	THE CRISIS	-1.65
26	DEFAULT	-1.63
27	BENEFITS	-1.56
28	UNEMPLOYED	-1.55
29	POVERTY	-1.52
30	SOCIAL SECURITY OFFICE	-1.51

Table 2. The FEARS index and KOSPI returns

This table links KOSPI index returns to the FEARS. The dependent variables include the following: contemporaneous returns, in Column (1); future returns in the next five weeks, in Columns (2), (3), (5), (6), and (7), respectively; and future KOSPI index returns over the first two weeks, in Column (4). The independent variable is the FEARS index and a set of control variables, which includes lagged KOSPI returns for up to five lags, the KOSPI200 volatility index (VKOSPI), and changes in the coincident composite index (CCI). T-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

$$\mathrm{Ret}_{\mathrm{i,t+k}} = \beta_0 + \beta_1 * \mathrm{FEARS}_{\mathrm{t}} + \sum_m r_m * Control^m_{\mathrm{i,t}} + \mathrm{u}_{\mathrm{i,t+k}}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t,t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
FEARS	-0.0566***	0.0064	-0.0212*	-0.0664***	0.0233**	-0.0049	-0.0211*
	(-5.04)	(0.54)	(-1.79)	(-3.51)	(1.99)	(-0.41)	(-1.78)
VKOSPI	-0.0382***	-0.0032	0.0146	-0.0311	0.0203	0.0218	0.0188
	(-2.87)	(-0.23)	(1.06)	(-1.38)	(1.47)	(1.57)	(1.36)
CCI	-0.0025	0.0000	0.0013	-0.0012***	0.0017	0.0019	0.0018
	(-0.78)	(0.01)	(0.40)	(-0.23)	(0.50)	(0.59)	(0.53)
Ret(t)		-0.0503	0.0120		-0.0579	-0.0051	-0.0380
		(-1.05)	(0.25)		(-1.21)	(-0.11)	(-0.80)
Ret(t-1)	-0.0738	0.0179	-0.0839*	-0.1227	-0.0045	-0.0192	0.0480
	(-1.63)	(0.38)	(-1.80)	(-1.61)	(-0.10)	(-0.41)	(1.03)
Ret(t-2)	0.0147	-0.0892*	-0.0032	-0.0597	-0.0224	0.0554	0.0523
	(0.32)	(-1.90)	(-0.07)	(-0.78)	(-0.48)	(1.18)	(1.12)
Ret(t-3)	-0.0980**	-0.0206	-0.0125	-0.1218	0.0431	0.0577	-0.0340
	(-2.16)	(-0.44)	(-0.27)	(-1.60)	(0.92)	(1.23)	(-0.73)
Ret(t-4)	-0.0045	-0.0217	0.0489	0.0244	0.0430	-0.0334	0.0340
	(-0.10)	(-0.46)	(1.04)	(0.32)	(0.92)	(-0.71)	(0.73)
Ret(t-5)	-0.0290	0.0341	0.0541	0.0509	-0.0433	0.0295	-0.1026**
	(-0.64)	(0.73)	(1.17)	(0.67)	(-0.94)	(0.63)	-2.2217
Constant	0.0099***	0.0011	-0.0035	0.0086	-0.0048	-0.0052	-0.0044
	(2.68)	(0.29)	(-0.90)	(1.37)	(-1.24)	(-1.35)	(-1.14)
Observation	469	468	467	467	466	465	464

Table 3. The FEARS index and other asset returns

This table reports the results of contemporaneous and predictive regressions of the FEARS and other index returns. Panel A includes the KOSPI200 equally- and value-weighted index returns. Panel B includes both large and small stock returns, and Panel C includes KOSDAQ and treasury-bond index returns. The dependent variables are contemporaneous returns, in Columns (1) and (4), and future returns. The independent variable is the FEARS index and a set of control variables (unreported), which includes: lagged returns for up to five lags, the KOSPI200 volatility index (VKOSPI), and changes in the coincident composite index (CCI). T-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

$$\mathrm{Ret}_{\mathrm{i,t+k}} = \beta_0 + \beta_1 * \mathrm{FEARS}_{\mathrm{t}} + \sum_m r_m * Control_{\mathrm{i,t}}^m + \mathrm{u}_{\mathrm{i,t+k}}$$

Panel A. KOSPI200 equally-weighted and value-weighted index returns

	KOSPI200 VW Index Returns			KOSPI200 EW Index Returns		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ret(t)	Ret(t,t+2)	Ret(t+3)	Ret(t)	Ret(t,t+2)	Ret(t+3)
FEARS	-0.0592***	-0.0686***	0.0262**	-0.0407***	-0.0535***	0.0180*
	(-5.14)	(-3.56)	(2.16)	(-3.79)	(-2.65)	(1.68)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	385	383	382

Panel B. Large stock returns and small stock returns

	Large stock returns			Small stock returns		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ret(t)	Ret(t,t+2)	Ret(t+3)	Ret(t)	Ret(t,t+2)	Ret(t+3)
FEARS	-0.0610***	-0.0701***	0.0265**	-0.0412***	-0.0554***	0.0065
	(-5.29)	(-3.64)	(2.17)	(-3.75)	(-2.81)	(0.57)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	469	467	466

Panel C. KOSDAQ returns and Treasury-bond index returns

	KOSDAQ			Treasury Index returns		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ret(t)	Ret(t,t+2)	Ret(t+3)	Ret(t)	Ret(t,t+2)	Ret(t+3)
FEARS	-0.0383***	-0.0452**	0.0024	0.0015	0.0039*	-0.0004
	(-3.00)	(-2.11)	(0.18)	(1.23)	(1.89)	(-0.36)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	469	467	466

Table 4. The FEARS index and limits to arbitrage

This table relates the FEARS index to weekly high-minus-low return spreads on portfolios sorted by either the CAPM beta, the volatility, the downside beta, or the downside volatility which are stock characteristics related to limits to arbitrage. The dependent variables are contemporaneous returns, in Columns (1), (4), (7) and (10), and future returns. The independent variable is the FEARS index and a set of control variables (unreported), which includes lagged returns for up to five lags, the KOSPI200 volatility index (VKOSPI), and changes in the coincident composite index (CCI). T-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Beta-sorted portfolio spread			Volatility-sorted portfolio spread		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ret(t)	Ret(t,t+2)	Ret(t+3)	Ret(t)	Ret(t,t+2)	Ret(t+3)
FEARS	-0.0386**	-0.0782**	0.0473**	-0.0307**	-0.0430*	0.0274*
	(-2.01)	(-2.52)	(2.48)	(-2.16)	(-1.84)	(1.90)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	469	467	466
	Downside Beta-sorted portfolio spreads			Downside Volatility-sorted portfolio spreads		
	(7)	(8)	(9)	(10)	(11)	(12)
	Ret(t)	Ret(t,t+2)	Ret(t+3)	Ret(t)	Ret(t,t+2)	Ret(t+3)
FEARS	-0.0114	-0.0550*	0.0323*	-0.0249*	-0.0388*	0.0309**
	(-0.62)	(-1.84)	(1.78)	(-1.89)	(-1.78)	(2.30)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	469	467	466

Table 5. The FEARS index and volatility

This table relates the FEARS index to market volatility. Panel A displays the seasonal-adjusted log realized volatility, which is calculated using KOSPI daily index returns. We then compute log-realized volatility, *rv*, and remove potential seasonal effects by regressing this on month-of-the year dummies. Panel B relays the log VKOSPI weekly index. Panel C tests total volatility, which is calculated by the standard deviation of individual stocks' daily returns in the KOSPI market. The independent variable is the FEARS index and a set of control variables (unreported), which include changes in the coincident composite index (CCI).

$$\mathrm{Vol}_{\mathrm{i,t+k}} = \beta_0 + \beta_1 * \mathrm{FEARS}_{\mathrm{t}} + \sum_m r_m * Control^m_{\mathrm{i,t}} + \mathrm{u}_{\mathrm{i,t+k}}$$

Panel D tests VKOSPI future returns against the FEARS index by Equation (3). T-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Adjusted realized volatility

	(1)	(2)	(3)	(4)	(5)
	Vol(t)	Vol(t+1)	Vol(t+2)	Vol(t+3)	Vol(t+4)
FEARS	0.0763*	-0.0492	-0.0452	-0.0035	-0.0017
	(1.66)	(-1.06)	(-0.98)	(-0.07)	(-0.04)
Controls	YES	YES	YES	YES	YES
Observation	469	468	467	466	465

Panel B. VKOSPI

	(1)	(2)	(3)	(4)	(5)
	Vol(t)	Vol(t+1)	Vol(t+2)	Vol(t+3)	Vol(t+4)
FEARS	0.0811**	0.0026	-0.0122	0.0217	0.0323
	(2.46)	(0.08)	(-0.37)	(0.65)	(0.98)
Controls	YES	YES	YES	YES	YES
Observation	469	468	467	466	465

Panel C. Total return volatility

	(1)	(2)	(3)	(4)	(5)
	Vol(t)	Vol(t+1)	Vol(t+2)	Vol(t+3)	Vol(t+4)
FEARS	0.1105**	0.0724	0.0332	0.0297	0.0279
	(2.40)	(1.56)	(0.71)	(0.64)	(0.61)
Controls	YES	YES	YES	YES	YES
Observation	469	468	467	466	465

Panel D. VKOSPI futures contract returns

	(1)	(2)	(3)	(4)	(5)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)
FEARS	0.2198**	0.1547	-0.0837	0.0641	0.0439
	(2.28)	(1.56)	(-0.83)	(0.62)	(0.43)
Controls	YES	YES	YES	YES	YES
Observation	83	82	81	80	79

Table 6. The FEARS index and fund flows

This table links the FEARS index to three fund groups specializing in equity (Panel A), bonds (Panel B), and MMF (Panel C). Fund flows are calculated for each group by the following regression:

$$\text{Flow}_{j,t} = \ln(\text{NAV}_{j,t}) - \ln(\text{NAV}_{j,t-1})$$

The independent variable is the FEARS index and a set of control variables (unreported), which includes lagged market returns for up to five lags, the KOSPI200 volatility index (VKOSPI), and changes in the coincident composite index (CCI). T-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

$$\text{Flow}_{I,t+k} = \beta_0 + \beta_1 * \text{FEARS}_t + \sum_m r_m * \text{Control}^m_{I,t} + u_{I,t+k}$$

Panel A. Equity fund flows

	(1) Ret(t)	(2) Ret(t+1)	(3) Ret(t+2)	(4) Ret(t+3)	(5) Ret(t+4)
FEARS	-0.0573***	-0.0029	-0.0290**	0.0264**	-0.0176
	(-4.54)	(-0.22)	(-2.29)	(2.11)	(-1.39)
Controls	YES	YES	YES	YES	YES
Observation	469	468	467	466	465

Panel B. Bond fund flows

	(1) Ret(t)	(2) Ret(t+1)	(3) Ret(t+2)	(4) Ret(t+3)	(5) Ret(t+4)
FEARS	-0.0056	-0.0082	-0.0126**	0.0057	-0.0151**
	(-0.89)	(-1.29)	(-1.99)	(0.88)	(-2.36)
Controls	YES	YES	YES	YES	YES
Observation	469	468	467	466	465

Panel C. MMF fund flows

	(1) Ret(t)	(2) Ret(t+1)	(3) Ret(t+2)	(4) Ret(t+3)	(5) Ret(t+4)
FEARS	0.0671***	0.0578**	-0.0330	-0.0546**	0.0088
	(2.97)	(2.55)	(-1.45)	(-2.42)	(0.38)
Controls	YES	YES	YES	YES	YES
Observation	469	468	467	466	465

Table 7. The FEARS index and trading behavior

This table relates the FEARS index to the Net Purchase Ratios (NPR) of three investor types: individual (Panel A), institutional (Panel B), and foreign (Panel C). The NPR is calculated for each group by the following regression:

$$\mathrm{NPR}_{\mathrm{i,t}} = (\mathrm{Buy}_{\mathrm{i,t}} - Sell_{i,t})/(\mathrm{Buy}_{\mathrm{i,t}} + Sell_{i,t})$$

The independent variable is the FEARS index and a set of control variables (unreported), which includes lagged market returns for up to five lags, the KOSPI200 volatility index (VKOSPI), and changes in the coincident composite index (CCI). T-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

$$\mathrm{NPR}_{\mathrm{i,t+k}} = \beta_0 + \beta_1 * \mathrm{FEARS}_{\mathrm{t}} + \sum_{\mathrm{m}} \mathrm{r}_{\mathrm{m}} * \mathrm{Control}^{\mathrm{m}}_{\mathrm{i,t}} + \mathrm{u}_{\mathrm{i,t+k}}$$

Panel A. Individual investors								
	KOSPI				KOSDAQ			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t+3)
FEARS	0.0301**	-0.0053	0.0206*	-0.0282**	-0.0045**	0.0002	-0.0020	-0.0006
	(2.58)	(-0.45)	(1.75)	(-2.38)	(-2.24)	(0.09)	(-0.97)	(-0.31)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observation	469	468	467	466	469	468	467	466
Panel B. Institutional investors								
	KOSPI				KOSDAQ			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t+3)
FEARS	0.0244	0.0161	0.0159	0.0492**	0.0358	0.0477	0.0549	-0.0104
	(1.12)	(0.71)	(0.68)	(2.09)	(1.00)	(1.29)	(1.48)	(-0.28)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observation	469	468	467	466	469	468	467	466
Panel C. Foreigner investors								
	KOSPI				KOSDAQ			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t)	Ret(t+1)	Ret(t+2)	Ret(t+3)
FEARS	-0.0781***	-0.0063	-0.0670**	0.0210	0.1124*	0.1193	0.0526	-0.0306
	(-2.79)	(-0.21)	(-2.25)	(0.70)	(1.79)	(1.88)	(0.83)	(-0.48)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observation	469	468	467	466	469	468	467	466

Table 8. Robustness checks

Table 8 displays the results from various robustness checks. We construct the FEARS index without winsorization and seasonalization in Panel A (Columns 1 to 3), and rebuild the FEARS index every three months (Columns 4 to 6). Panel B considers additional controls, including the KOSPI's top dummy (Columns 1 to 3) and holiday controls (Columns 4 to 6). Panel C considers robustness, including estimates when the top 25 terms are used (Columns 1 to 3), and the top 35 terms are used (Columns 4 to 6). The independent variable is the FEARS index and a set of control variables (unreported), which includes lagged KOSPI returns for up to five lags, the KOSPI200 volatility index (VKOSPI), and changes in the coincident composite index (CCI). T-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. FEARS index with different construction

	without winsorization and seasonalization			Rebuild every 3 month		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ret(t)	Ret(t,t+2)	Ret(t+3)	Ret(t)	Ret(t,t+2)	Ret(t+3)
FEARS	-0.0602 ***	-0.0708 ***	0.0259 **	-0.0569 ***	-0.0640 ***	0.0288 **
	(-5.56)	(-3.88)	(2.25)	(-4.46)	(-2.97)	(2.11)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	469	467	466

Panel B. FEARS index with additional controls

	KOSPI Top dummy			Holiday dummy		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ret(t)	Ret(t,t+2)	Ret(t+3)	Ret(t)	Ret(t,t+2)	Ret(t+4)
FEARS	-0.0498***	-0.0618***	0.0243*	-0.0663***	-0.0730***	0.0284**
	(-3.49)	(-2.67)	(1.73)	(-5.03)	(-3.33)	(2.05)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	469	467	466

Panel C. FEARS index with the top 25 terms and the top 35 terms

	Top 25			Top 35		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ret(t)	Ret(t,t+2)	Ret(t+4)	Ret(t)	Ret(t,t+2)	Ret(t+3)
FEARS	-0.0639***	-0.0668***	0.0216*	-0.0522***	-0.0601***	0.0243*
	(-5.28)	(-3.26)	(1.68)	(-4.01)	(-2.76)	(1.81)
Controls	YES	YES	YES	YES	YES	YES
Observation	469	467	466	469	467	466