Available online at www.sciencedirect.com

Borsa İstanbul Review

Borsa İstanbul Review 22-2 (2022) 226–239
http://www.elsevier.com/journals/borsa-istanbul-review/2214-8450

Full Length Article

Does a search attention index explain portfolio returns in India?

Munusamy Dharani [a], M. Kabir Hassan [b,*], Mohammad Zoynul Abedin [c], Mohd Adib Ismail [d]

[a] Department of Finance and Accounting, Indian Institute of Management Kashipur, Uttarakhand −244713, India
[b] Department of Economics and Finance, University of New Orleans, New Orleans, LA 70148, USA
[c] Department of Finance and Banking, Hajee Mohammad Danesh Science and Technology University, Dinajpur 5200, Bangladesh
[d] Universiti Kebangsaan Malaysia, 43600, UKM Bangi, Selangor, Malaysia

Received 29 March 2020; revised 11 April 2021; accepted 13 April 2021
Available online 22 April 2021

Abstract

Employing asset-pricing models over the period 2012 to 2017, this study examines whether a search attention index (SAI) explains the variation in the weekly excess return of stocks. The study finds that the estimated abnormal return of a portfolio based on search intensity is significantly high for stocks with higher search intensity and low for stocks with lower search intensity. Further, the study observes that, when the SAI is high, the excess returns are high for stocks with a high value, high volatility, and high sensitivity. Interestingly, the study documents that in the Indian market investor attention is irrelevant for stocks with extremely high risk. This study finds that the SAI in India explains the variation in the excess return of stocks as well as the market, size, value, and momentum factors.

JEL classification: G11; G12; G14

Keywords: Asset pricing; Google search volume index; Investor sentiment; Stock market

1. Introduction

Fama (1970) argues that present market value reflects all available market information. He shows that investors are rational and use cognitive information when picking stocks in a rational market. Further, he assumes that investors cannot earn an abnormal return in the market. Therefore, researchers expect the asset-pricing models to have an insignificant intercept that is expected to be zero. In general, the share price is determined by the risk and return characteristics of a firm. Early studies document that market factors explain variations in the excess returns of the cross-section of stocks. However, market factors are inconsistent in explaining the variations in stock returns (Banz, 1981; Chan et al., 1991; Fama and French, 1992; Stattman, 1980). To clarify this issue, the seminal study by Fama and French (1993) recommends that the variation in stock returns by determined by the market premium, size, and firm value. Additionally, Jagadeesh and Titman (1993) identify momentum behavior: the tendency of stocks that perform well in a prior year to continue this performance in the next year. Taking all this into account, Carhart (1997) proposes a four-factor model for explaining the variation in stock returns. Then, Fama and French (2015) propose a five-factor model that extends the three-factor model by incorporating profitability and investment to explain the variation in stock returns. However, studies on variations in stock returns are not limited to these factors. All asset-pricing models argue that investors are rational and have infinite cognitive resources to access the information available in the market. However, in the real world, investors collect a limited set of information because of the constraints of time and labor. In other words, investors are irrational and have finite

* Corresponding author.
E-mail addresses: dharani@iimkashipur.ac.in (M. Dharani), mhassan@uno.edu (M.K. Hassan), abedinmz@yahoo.com (M.Z. Abedin), mohadis@ukm.edu.my (M.A. Ismail).
Peer review under responsibility of Borsa İstanbul Anonim Şirketi.

https://doi.org/10.1016/j.bir.2021.04.003

cognitive ability (Barber & Odean, 2008; Kahneman, 1973). Information is a highly valuable resource for determining investor behavior and the effect of their attention on asset prices.

As a result, Merton (1987) argues that investor attention is also one of the factors that explains stock returns. The literature documents that the prevailing asset-pricing models do not consider the effect of investor attention in determining variations in stock returns (Baker & Wurgler, 2006; Barber & Odean, 2008; Bijl et al., 2016; Chen et al., 2013; Dowling & Lucey, 2005; Da et al., 2011; Edmans et al., 2007; Hirshleifer and Shumway (2003); Joseph et al., 2011; Kaplanski & Levy, 2010; Palomino et al., 2009). However, Barber and Odean (2008), Baker and Wurgler (2006), Da et al. (2011), Joseph et al. (2011), Peng and Xiong (2006), and Sims (2003) empirically document that investor attention is actually an essential factor in explaining variations in stock returns. The evolving literature measures investor attention in the market using proxies for advertising expenses (Gustavo et al., 2004), new information (Baker & Wurgler, 2006), news articles (Tetlock, 2007), extreme returns (Barber & Odean, 2008), trading volume (Barber & Odean, 2008), and news and headlines (Yuan, 2008). More recently, Da et al. (2011) empirically examine a novel and direct proxy for measuring investor attention in the market using Google search intensity via Google Trends.

The growing literature argues that investor sentiment measured by the Google search intensity is positively associated with the returns and trading volume of stocks (Adachi et al., 2017; Bank et al., 2011; Da et al., 2011; Joseph et al., 2011; Taketa & Wakao, 2014; Vlastakis & Markellos, 2012). However, Chen (2017) reports a negative relationship between investor attention and index returns. Further, Chen documents that high investor attention predicts higher returns in the US stock market but not in other markets. Bijl et al. (2016), Da et al. (2011), Joseph et al. (2011), and Yung and Nafar (2017) claim that when Google search intensity is high, the stock return is also high. Google search intensity also predicts the volatility of the equity and commodity markets (Afkhami et al., 2017; Smith, 2012). Moreover, Kim et al. (2018) argue that an increase in Google search intensity leads to an increase in the volatility and trading volume but not abnormal returns. Swamy and Dharani (2019) and Swamy et al. (2019) find that when Google search intensity is high, the stock return is also high. They find a positive relationship between Google search intensity and stock returns in India. In analyzing previous studies, we find mixed results on the Google search volume and stock returns in developed and developing economies.

In addition, most studies use the raw search volume index as a proxy for the investor attention variable. The present study constructs a search attention index (SAI) and examines whether it explains variations in the excess returns of stocks using asset-pricing models in emerging markets, such as India. In particular, the study considers stocks in the S&P Bombay Stock Exchange (BSE) 500 index, which comprises 500 stocks and covers nearly 93 percent of the market capitalization of the BSE, one of the oldest stock exchange in the world. The BSE is the tenth-largest stock exchange, with a market capitalization of more than $1.2 trillion (World Federation of Exchanges, 2018). Further, on average, 83 percent of retail investors invest directly in stocks (*Economic Times* report, July 31, 2017). Participation by individual investors in equity investment increased 30 percent over the period 2017–2018 (*Economic Times* report, December 10, 2018), which shows that retail investors are playing an increasingly important role in the Indian stock market. In general, retail investors try to get information about the stocks before buying them. They use a search engine to find information on the past behavior and financial details of the stocks. Therefore, when they use a search engine to seek stock details, the search intensity of the stock gradually increases. This indicates that the investors intend to buy the stocks, and then the prices of the stocks increase in the short run. Thus, investors seek to obtain additional returns in the short run based on information asymmetry in the market. These scenarios lead us to examine whether investor search intensity explains variations in stock returns based on the Indian stock market.

The remainder of the paper is arranged as follows. Section 2 highlights the previous studies on Google search intensity and its effects on stock returns, volatility, and trading volume. Section 3 describes the data and methodology. Section 4 discusses the empirical results and implications. Concluding remarks are provided in Section 5.

2. Literature review

Since the 2010s, the interrelation between Google search intensity and stock returns has been a topic of great curiosity. An innovative paper by Da et al. (2011) argues that the Google search volume index (GSVI) actively leads alternative direct proxies for measuring investor attention and capturing retail investor behavior in the US. Further, they find that the GSVI predicts higher stock returns in the following two weeks and price reversal within one year. Using the Carhart (1997) four-factor model, Joseph et al. (2011) investigate the impact of Google search intensity on the returns and trading volume of stocks over the period 2005 to 2008. Joseph et al. find that Google search intensity predicts the abnormal returns and trading volume of the stocks. Similarly, Bank et al. (2011) examine the influence of Google search intensity on stock market activity from 2004 to 2010. They find that Google search intensity is significantly associated with stock returns and liquidity in the German stock market. Taketa and Wakao (2014) investigate the influence of Google search volume (GSV) on the returns and trading volume of stocks in the Japanese stock market using a sample of 189 stocks from 2008 to 2011. Their study documents that GSV strongly affects the trading volume and weakly influences stock returns in the Japanese market. Mnif et al. (2020) use three social media databases and show that these sentiment measures have a remarkable impact on contemporaneous and lagged returns of various Islamic assets. Metawa, Hassan, and Elhoseny (2017) use an intelligent model based on a genetic algorithm (GA) to organize bank lending decisions in a highly competitive environment with a credit crunch constraint. Abedin et al. (2019) use 12 feature selection methods for support vector machine (SVM) classifiers, checking their optimality by comparing them to some statistical and baseline methods with Chinese data.

Interestingly, Da et al. (2015) construct a new measure of investor sentiment, such as the Financial and Economic Attitudes Revealed by Search (FEARS) index and find that it predicts short-term return reversals and temporary increases in stock volatility over the period 2004 to 2011. Similarly, Siganos (2013) examines the impact of the GSVI on stock price prediction using a sample of 430 stocks over the period 2004−2010, finding that the GSVI explains variations in stock prices. Yung and Nafar (2017) examine the influence of GSV as a proxy for measuring retail investor attention on the returns of real estate investment trusts (REIT) over the period 2004 to 2012. They argue that when GSV is higher, the REITs have higher expected returns. Further, over time, an increase in returns is followed by a reversal. Using a cross-sectional regression model, Ying et al. (2015) document that GSV affects stock market returns positively in the Chinese stock market. By contrast, Bijl et al. (2016) investigate the impact of Google Trends on stock predictions for a sample of 500 companies in the US on the S&P 500 index over the period 2008 to 2009, finding that a high GSVI negatively influences stock returns.

Google search also affects stock volatility. For example, using a GARCH (1,1) model, Smith (2012) argues that GSV predicts volatility in foreign currency. Afkhami et al. (2017) examine the forecasting ability of GSV on six commodity prices using GARCH models. They find that GSV is a significant predictor of volatility in energy markets over the period 2004 to 2016. Kim et al. (2018) investigate the predictive power of GSV on the returns, volatility, and trading volume of stocks from 2012 to 2017. They find that an increase in GSV helps to predict the volatility and trading volume, but not abnormal returns, of stocks on the Oslo Stock Exchange. Likewise, Vlastakis and Markellos (2012) argue that demand for information, measured by GSV, is positively associated with volatility and the trading volume of major stocks traded on the NYSE and the NASDAQ. In the same way, Tantaopas et al. (2016) examine the effect of GSV on the returns, volatility, and trading volume of stocks in developing markets. Their study documents that changes in GSV significantly affect the returns, volatility, and trading volume of the stocks. Further, to measure the behavior of institutional investors, Ben-Rephael et al. (2017) suggest a direct measurement of abnormal institutional investor attention (AIA) based on the frequency of searching for and reading news about the stocks on Bloomberg terminals. They find that institutional attention reacts rapidly to major news events, leads retail attention, and help achieve permanent price adjustment.

Das and Ziobrowski (2015) document that the online search indices in India are significantly related to future movement in real estate stocks. Similarly, Venkataraman et al. (2018) observe that Google search intensity predicts housing prices in India. Further, Swamy and Dharani (2019) investigate whether the GSVI predicts stock returns of the Nifty 50 companies, employing panel data from July 2012 to June 2017. Their study finds that a high GSVI leads to positive returns in the Indian stock market. Then Swamy et al. (2019) examine the impact of the GSVI on the stock returns of the S&P BSE 500 companies using a quantile regression model over the period 2012 to 2019. They find that a higher GSVI predicts positive and significant returns in the first and second weeks. Finally, the study supports the findings of a cointegration relationship between the GSVI and stock returns in India.

Most of these studies consider GSV as one of the explanatory variables in the model to examine its impact on stock returns, volatility, and trading volume. In this study, we first construct the SAI using data on GSV. Then, we examine whether the SAI explains variations in the excess returns of stocks using a sample of 436 companies in India. To the best of our information, no prior study has investigated whether the SAI explains stock returns in India.

3. Data and methodology

3.1. Sample selection

In this study, we consider sample companies on the S&P BSE 500 index, which comprises 500 highly traded companies in India. We use a list of stock prices, market capitalization, price-to-book value, the online search intensity index from Google Trends, stock beta, and 12-week stock volatility of the companies. The financial and accounting data for the stocks come from the Prowess database of the Centre for Monitoring Indian Economy (CMIE) for the period August 2012 to July 2017. We retrieve weekly online search intensity data for every stock from https://trends.google.com/trends/?geo=IN/for the same period. When we collected the Google search intensity data in 2017, we identified monthly data from 2004 onward, but Google Trends provides weekly data for the preceding five years from the date of collection. Based on Da et al. (2011), Joseph et al. (2011), and others, we decided to consider weekly data for the five years from 2012 to 2017. We used the ticker name to find the Google search intensity in Google Trends. Da et al. (2011) explain that GSV captures the sentiment behavior of retail investors and is a proxy for the sentiment index in the US market. The financial, accounting, and search intensity dataset for 64 companies is incomplete for the sample period. Therefore, we use a sample size of 436 stocks for the entire analysis. Further, the study uses factors such as SMB (small minus big), HML (high minus low), WML (winning minus losing), and the Treasury bill rate as a proxy for the risk-free rate (Rf) (Agarwalla et al., 2013).

3.2. Portfolio formation

We form portfolios from the 436 stocks based on the approach used by Fama and French (1993). Initially, at the beginning of every week, we sort the stocks by the weekly

GSV and form five quintiles from Q1 to Q5, in which Q1 includes the stocks with the lowest Google search intensity and Q5 includes the stocks with the highest Google search intensity. Then, we create the SAI, which is the excess return of Q5 over Q1 (Joseph et al., 2011). In other words, the difference in returns between stocks with high search intensity (Q5) and stocks with low search intensity (Q1) is the search attention index (Q5-Q1). Further, for every week from September 2012 to July 2017, we divide all companies into five equally populated groups based on market capitalization, price-to-book value, beta, and long-run volatility. The portfolios are unchanged for the week and then resorted every week. Each quantile is made up of 87 stocks. Each stock has 257 weekly returns, and each quantile comprises weekly Google search intensity for 257 stocks. Then, we form the weekly time-series return for each quantile by taking an average of 257 stocks. Additionally, double-sorted portfolios are formed by size and value. Finally, following the literature, the 30 single-sorted, 10 double-sorted, and 7 long-short portfolios are analyzed with asset-pricing models.

3.3. Methodology

We initially employ the capital asset-pricing model (CAPM) proposed by Sharpe-Linter-Mossin to estimate the expected returns on the stocks.

$$R_{pt} - R_{ft} = \alpha + \beta_m \left(R_{mt} - R_{ft}\right) + \varepsilon_t \tag{1}$$

where R_{pt}-R_{ft} is the excess portfolio return, and $(R_m - R_f)$ is the excess market return. α and β_m are the estimated coefficients. The CAPM model implies that the excess market returns fully explain the portfolio's risk-adjusted outcomes. In this model, the estimated coefficient of the intercept (α) is expected to equal zero. However, a significantly positive or negative intercept implies that a portfolio yields a higher abnormal return or lower abnormal return in the market. Further, it also explains the overperformance or underperformance of the portfolio. A majority of the literature documents excess returns beyond the market risk factor. In other words, the CAPM is not a reliable model for explaining excess portfolio returns (Banz, 1981; Chan et al., 1991; Stattman, 1980).

As a result, Fama and French (1992, 1993) argue that market capitalization (a proxy for the size factor) and the book-to-market ratio (a proxy for the value factor) are also important factors in explaining variations in stock returns. Since their work, these two factors have become the most widely used explanatory variables in the asset-pricing models in developed and developing markets. Fama and French's (1993) three-factor model augments the CAPM with two additional factors to capture the size and value premiums of the portfolios:

$$R_{pt} - R_{ft} = \alpha + \beta_m \left(R_{mt} - R_{ft}\right) + \beta_S SMB_t + \beta_h HML_t + \varepsilon_t \tag{2}$$

where R_{pt}-R_{ft} is the excess return on the test portfolio; R_f is the risk-free rate; and $(R_m - R_f)$ is the excess market return. SMB is the difference in returns between stocks with the lowest market capitalization and those with the highest market capitalization. HML is the difference in returns between stocks with the highest price-to-book value and those with the lowest price-to-book value. Jagadeesh and Titman (1993) document that stock returns also exhibit momentum behavior in the US market. They argue that stocks that perform well for the prior year tend to continue to do well. As a result, Carhart (1997) adds a momentum factor to the Fama and French (1993) three-factor model and proposes a four-factor model to capture the effect of momentum on returns. The Carhart (1997) four-factor model is:

$$\begin{aligned} R_{pt} - R_{ft} = {} & \alpha + \beta_m \left(R_{mt} - R_{ft}\right) + \beta_S SMB_t + \beta_h HML_t \\ & + \beta_u WML_t + \varepsilon_t \end{aligned} \tag{3}$$

where WML is the difference in returns between stocks with the highest returns in the prior year and those with the lowest returns in the prior year. Baker and Wurgler (2006), Barber and Odean (2008), Da et al. (2011), and Merton (1987) show that investor sentiment also explains stock returns. Joseph et al. (2011) propose a new sentiment index based on GSV, which is the difference in returns between stocks with the highest search intensity and those with the lowest search intensity. We create the SAI, a proxy for the sentiment factor, based on GSV and add it to the asset-pricing models as an explanatory variable in the context of an emerging market such as India. Finally, we add the SAI as an explanatory variable to investigate whether it explains variations in the excess returns of different portfolios. Accordingly, we estimate the models as:

$$ER_t = \alpha + \beta_{mkt}\left(R_{m,t} - R_{f,t}\right) + \beta_{sai} SAI_t + \varepsilon_t \tag{4}$$

$$\begin{aligned} ER_t = {} & \alpha + \beta_{mkt}\left(R_{mt} - R_{ft}\right) + \beta_{smb} SMB_t + \beta_{hml} HML_t \\ & + \beta_{sai} SAI_t + \varepsilon_t \end{aligned} \tag{5}$$

$$\begin{aligned} ER_t = {} & \alpha + \beta_{mkt}\left(R_{mt} - R_{ft}\right) + \beta_{smb} SMB_t + \beta_{hml} HML_t + \beta_{wml} WML_t \\ & + \beta_{sai} SAI_t + \varepsilon_t \end{aligned} \tag{6}$$

Merton (1987) argues that stocks with low investor recognition have to offer higher returns to compensate their holders for being imperfectly diversified. Baker and Wurgler (2006) state that when investor sentiment is low, returns are relatively high for stocks with low market capitalization. Barber and Odean (2008) claim that when investor attention increases, returns are high in the short run. Therefore, stocks that capture investor attention tend to generate excess returns and high trading volumes temporarily. Moreover, because of the growth of the internet, many retail investors search for stock information using internet search engines such as Google and Yahoo. Therefore, we expect the SAI to have a significant effect in explaining variations in the excess returns of different portfolios.

4. Empirical results and discussion

4.1. Summary statistics

Table S1 (see Table S1, available online) describes the variables used in the study, and Table 1 presents the summary statistics of the 30 test portfolios and explanatory variables. The results indicate that the average return in Q1 is 0.144, with a standard deviation of 2.411, whereas the average return in Q5 is 0.608, with a standard deviation of 2.567. This clearly shows that when Google search intensity is low, the average stock return is low. Stock returns gradually increase when Google search intensity increases.

Then, to examine the size effect, we divide the stocks into five quantiles from S1 to S5 by the market capitalization at the beginning of every week. S1 represents stocks with the lowest market capitalization, and S5 represents stocks with the highest market capitalization. Because the stocks are sorted weekly by Google search intensity, we perform the same exercise for the other variables as well. Agarwalla et al. (2013) report that 90 percent of the stocks in the Indian capital market have small market capitalization. The results also reveal that the average return is higher on stocks with the lowest market capitalization than those with the highest market capitalization.

Furthermore, we divide the stocks into five quantiles from V1 to V5 by the price-to-book value. The average return on stocks with a low value is −0.052, with a standard deviation of 3.443, whereas the average return on those with a high value is 0.573, with a standard deviation of 1.854. This clearly shows that high-value stocks provide higher returns than low-value stocks. In addition, the average return on small-cap stocks with a high value is 0.717, with a standard deviation of 2.316, whereas the average return on big-cap stocks with a high value is 0.493, with a standard deviation of 1.816. This further demonstrates that high-value stocks yield higher returns in the market. Furthermore, the average return on stocks with the highest beta is 0.201, with a standard deviation of 3.539. This

Table 1
Summary statistics (N = 257).

	Mean	Median	Max.	Min.	SD	Skewness	Kurtosis
Low search intensity Q1	0.144	0.326	9.611	−8.334	2.411	−0.192	4.453
Q2	0.243	0.430	8.471	−8.145	2.359	−0.257	4.137
Q3	0.311	0.635	10.142	−8.358	2.344	−0.255	4.928
Q4	0.371	0.500	8.409	−7.275	2.381	−0.268	3.992
High search intensity Q5	0.608	0.829	11.110	−7.717	2.567	−0.108	4.376
Small firms S1	0.352	0.640	12.719	−10.744	3.028	−0.231	4.956
S2	0.387	0.648	11.742	−8.779	2.568	−0.149	5.141
S3	0.359	0.553	9.448	−7.616	2.313	−0.187	4.215
S4	0.330	0.541	9.263	−7.211	2.340	−0.216	4.149
Large firms S5	0.267	0.355	6.151	−5.873	2.058	−0.086	3.145
Low-value firms V1	−0.052	0.157	14.703	−11.576	3.443	−0.120	4.902
V2	0.291	0.555	12.576	−8.209	2.657	−0.041	5.154
V3	0.392	0.569	9.023	−8.120	2.368	−0.215	4.110
V4	0.489	0.586	7.419	−7.080	2.052	−0.300	4.230
High-value firms V5	0.573	0.703	5.511	−5.750	1.854	−0.280	3.573
Small & low-value SV1	−0.077	0.167	15.589	−11.941	3.538	−0.094	4.961
SV2	0.324	0.603	13.837	−10.122	2.989	−0.133	5.127
SV3	0.396	0.571	10.257	−8.890	2.637	−0.296	4.552
SV4	0.484	0.540	9.617	−8.014	2.372	−0.184	4.300
Small & high-value SV5	0.717	0.953	8.358	−7.180	2.316	−0.163	4.307
Large and low-value BV1	−0.035	−0.067	13.194	−11.342	3.458	−0.027	4.463
BV2	0.267	0.337	10.302	−8.075	2.526	0.069	4.248
BV3	0.405	0.420	7.694	−7.299	2.149	−0.153	3.704
BV4	0.429	0.538	6.144	−6.431	1.812	−0.324	3.787
Large & high-value BV5	0.493	0.548	4.665	−4.711	1.816	−0.245	3.077
Low-beta B1	0.325	0.389	4.347	−5.417	1.577	−0.352	3.730
B2	0.359	0.426	7.885	−6.029	1.968	−0.180	4.089
B3	0.396	0.552	10.072	−8.210	2.401	−0.123	4.824
B4	0.407	0.649	11.930	−8.998	2.805	−0.189	4.695
High-beta B5	0.201	0.520	14.199	−11.666	3.539	−0.189	4.423
RF	0.030	0.030	0.044	0.011	0.005	−0.174	5.177
RM	0.073	0.094	1.296	−1.416	0.420	−0.084	3.600
RM_RF	0.043	0.066	1.263	−1.444	0.420	−0.087	3.603
SMB	0.032	0.044	1.447	−1.054	0.369	0.057	4.273
HML	0.022	0.014	3.019	−1.500	0.539	0.693	6.396
WML	0.055	0.116	1.272	−3.016	0.469	−1.397	10.249
SAI	0.464	0.458	3.734	−2.479	0.884	−0.069	3.642

shows that when stocks are much more sensitive than the market, the stock return is low and highly volatile. Investors face difficulty in arbitraging high-volatility stocks in the market (Baker & Wurgler, 2006; Joseph et al., 2011).

4.2. Correlation matrix

Table 2 reports the correlation matrix of the variables. The results show a highly significant and positive relationship between the SAI and stock returns, with a value of 17.3. The SAI is positively correlated with the market and momentum factors and negatively correlated with the size factor. Additionally, the stock return is positively correlated with the market, size, and value factors, but negatively correlated with the momentum factor. The overall results of Table 2 reveal that the SAI is a factor that should also be considered when examining variations in stock returns.

4.3. Google search intensity and returns from the portfolios

The asset-pricing models assume that all information is freely available to investors. However, because of time, workforce, technology, and other factors, investors obtain only limited information in the market, so they seek information using the Google search engine to obtain operational and financial details about the stocks. Therefore, Google search intensity plays an important role in measuring investor attention in the market. Therefore, we need to investigate information asymmetry in the market, which can be measured by the intercept in the models. To examine the ability of Google search intensity to forecast stock returns, for each portfolio sorted by the GSVI, we employ the CAPM, Fama and French three-factor model (1993), and Carhart four-factor model (1997). The results of Table 3 show that the estimated alpha coefficients for stocks with low search intensity are negative and highly significant in the three-factor and four-factor models. Further, the estimated coefficients for stocks with high search intensity are positive and highly significant. The adjusted R squared for each portfolio is from 80 percent to 91 percent. Fama and French (1993, 2012) state that the market, size, and value factors explain the cross-section of stock returns in global markets. The results of Table 3 further illustrate that when the level of search intensity increases, the abnormal returns associated with the corresponding portfolio also increase (Joseph et al., 2011). Moreover, the results show that the difference in returns between stocks with high search intensity and those with low intensity across the different estimation procedures is from 0.432 to 0.459 and is highly significant at the 1 percent level. In other words, a portfolio in the Indian stock market that is long on Q5 and short on Q1 has a weekly abnormal return of from 0.432 to 0.459. These results are consistent with the results of Joseph et al. (2011) in the US market. The overall results indicate that when retail investors search for stock details on the internet, their buying behavior increases in the short term. As a result, stock returns increase in the short run and then decline in the long run. This result supports the "buying pressure hypothesis" proposed by Barber and Odean (2008).

4.4. Size portfolios and the search attention index (SAI)

This part of the study examines whether the SAI explains the excess returns of the portfolios sorted by market cap. Banz (1981) argues that stocks with small market capitalization provide higher returns than stocks with large market capitalization. Fama and French (1992, 1993) document that the size factor measured by market capitalization explains the cross-section of stock returns in the US market. To test the size effect, we sort the stocks into five quantiles at the end of each week through S1 to S5 based on weekly market caps. Portfolio1 is a portfolio of stocks with low market caps and portfolio5 is a portfolio of stocks with high market caps. The weekly cross-section average returns are estimated for each portfolio, which consists of 87 stocks. The asset-pricing models are then employed over the period of September 2012 to July 2017. Table 4 (see Table S2, available online) reports that the estimated alpha coefficients are mostly insignificant and are from 0.005 to 0.35 under the Fama and French (1993) and Carhart (1997) models. This shows that the models capture variations in the returns explained by the selected factors. Further, the SAI coefficients are negatively insignificant for small-cap stocks and positively insignificant for large-cap stocks. Moreover, the market and size factors fully explain variations in stock returns when the market capitalization is extremely high. The market, size, and value factors explain variations in stock returns when the market capitalization is extremely low.

Table 2
Correlation matrix.

	RET	SAI	RM_RF	SMB	HML	WML
RET	1	0.173***	0.930***	0.136**	0.669***	−0.290***
SAI	0.173***	1	0.241***	0.269***	−0.022	0.118*
RM_RF	0.930***	0.241***	1	−0.096	0.555***	−0.261***
SMB	0.136**	−0.269***	−0.096	1	0.042	0.073
HML	0.669***	−0.022	0.555***	0.042	1	−0.40***
WML	−0.290***	0.118*	−0.261***	0.073	−0.40***	1

Note: *, **, and *** represent significance levels at 10%, 5%, and 1%, respectively. RET stands for the weekly returns. SAI is the search attention index. RM_RF is the weekly excess market returns. SMB is the small minus big stocks. HML is the high minus low stocks. WML is the winning minus and losing stocks. The variable definitions are reported in Table S1 (available online).

Table 3
Returns of the portfolios, sorted by Google Search Volume Index in every week ($N = 257$).

	Models	Intercept	RM_RF	SMB	HML	WML	Adj. R^2
Q1	CAPM	−0.078 (0.066)	5.166*** (0.157)				0.808
	FF	−0.122***	4.508***	1.519***	1.104***		0.915
		(0.044)	(0.127)	(0.120)	(0.099)		
	Carhart	−0.103**	4.492***	1.553***	1.002***	−0.312***	0.918
		(0.044)	(0.125)	(0.119)	(0.103)	(0.101)	
Q2	CAPM	0.024	5.103***				0.824
		(0.062)	(0.148)				
	FF	−0.018	4.611***	1.386***	0.854***		0.904
		(0.046)	(0.132)	(0.125)	(0.103)		
	Carhart	−0.020	4.613***	1.384***	0.862***	0.025	0.904
		(0.047)	(0.133)	(0.126)	(0.109)	(0.107)	
Q3	CAPM	0.093	5.073***				0.825
		(0.062)	(0.146)				
	FF	0.046	4.656***	1.517***	0.765***		0.910
		(0.044)	(0.127)	(0.120)	(0.099)		
	Carhart	0.043	4.658***	1.513***	0.776***	0.035	0.910
		(0.045)	(0.127)	(0.121)	(0.104)	(0.103)	
Q4	CAPM	0.147	5.216***				0.845
		(0.059)	(0.140)				
	FF	0.106**	4.824***	1.310***	0.705***		0.909
		(0.045)	(0.130)	(0.123)	(0.101)		
	Carhart	0.108**	4.823***	1.312***	0.698***	−0.021	0.909
		(0.046)	(0.130)	(0.124)	(0.107)	(0.105)	
Q5	CAPM	0.364***	5.673***				0.860
		(0.060)	(0.143)				
	FF	0.337***	5.194***	0.969***	0.788***		0.902
		(0.051)	(0.146)	(0.138)	(0.113)		
	Carhart	0.336***	5.194***	0.968***	0.792***	0.012	0.901
		(0.051)	(0.146)	(0.139)	(0.120)	(0.118)	
Q5-Q1	CAPM	0.442***	0.507***				0.054
		(0.054)	(0.128)				
	FF	0.459***	0.686***	−0.550***	−0.316***		0.134
		(0.052)	(0.149)	(0.141)	(0.115)		
	Carhart	0.439***	0.702***	−0.585***	−0.210*	0.324***	0.156
		(0.052)	(0.147)	(0.139)	(0.120)	(0.119)	

Note: *, **, and *** represent significance levels at 10%, 5%, and 1%, respectively. RET stands for the weekly returns. SAI is the search attention index. RM_RF is the weekly excess market returns. SMB is the small minus big stocks. HML is the high minus low stocks. WML is the winning minus and losing stocks. The variable definitions are reported in Table S1 (available online). The standard errors are reported in the parentheses. CAPM is the capital asset, pricing model. FF stands for Fama and French Model. Carhart is the Carhart four-factor model.

The primary purpose of this study is to examine whether the SAI explains variations in stock returns in India. When the coefficients of the SAI increase from portfolio1 to portfolio4, the alpha coefficients decrease. This shows a negative relationship between the SAI and excess returns on stocks in the short run. Further, the SAI coefficients are highly significant for stocks with market capitalization is from 20 percent to 80 percent. Agarwalla et al. (2013) report that 90 percent of the stocks in the Indian stock market have low market capitalization. Therefore, the results confirm that the SAI explains variations in the excess returns of stocks based on market capitalization. The overall results reveal that when search attention is high, excess returns on stocks are low, and vice-versa. These results support the investor recognition hypothesis proposed by Merton (1987), which states that stocks with low investor recognition have to offer higher returns to compensate their owners for being imperfectly diversified. The reason for this is that investors, who have incomplete information, are not aware of all the securities in the market.

4.5. Value portfolios and the search attention index (SAI)

In this section, we examine the value effect on the returns of portfolios based on their price-to-book value. Stattman (1980) argues that stocks with a higher value earn high returns in the market. Fama and French (1992) also support the claim that the value effect exists in the US market, and propose a model (1993) to capture the size and value effect in addition to the market factor. To examine the value effect, we divide the stocks into five quantiles from portfolio1 to portfolio5 based on their price-to-book value. Porfolio1 is a portfolio of stocks with a low price-to-book value (value stocks), and portfolio5 is a portfolio of stocks with a high price-to-book value (growth stocks). For each portfolio, we estimate the time-series returns from September 2012 to July 2017, consisting of 257 observations. Then, we estimate the asset-pricing models for each portfolio sorted by value. The results of the estimated models are presented in Table 4 (see Table S3, available online). The alpha coefficients are negative

Table 4
Returns of the portfolios sorted by the market capitalization, price to book value (PB), small & price-to-book value (PB), large & price-to-book value, the beta, and the 12-week standard deviation of the company (N = 257).

	Models	Market capitalization		Price-to-book value		Small and PB		Large and PB		Beta		Volatility	
		Intercept	SAI	Intercept	SAI	Intercept	SAI	Intercept	SAI	Intercept	SAI	Intercept	SAI
Portfolio 1	CAPM	0.350***	−0.577***	−0.206*	−0.352***	−0.136	−0.557***	−0.303***	−0.101	0.204***	−0.038	−0.005	0.022
		(0.115)	(0.119)	(0.115)	(0.119)	(0.131)	(0.135)	(0.115)	(0.119)	(0.058)	(0.060)	(0.039)	(0.040)
	FF	0.040	−0.090	−0.396***	0.028	−0.425***	−0.043	−0.344***	0.067	0.113**	0.081	−0.071**	0.112***
		(0.070)	(0.074)	(0.080)	(0.085)	(0.084)	(0.089)	(0.097)	(0.103)	(0.052)	(0.056)	(0.034)	(0.036)
	Carhart	0.040	−0.090	−0.378***	0.108	−0.410***	0.022	−0.318***	0.183*	0.106**	0.047	−0.074**	0.099***
		(0.070)	(0.075)	(0.075)	(0.081)	(0.081)	(0.087)	(0.089)	(0.095)	(0.051)	(0.055)	(0.034)	(0.036)
Portfolio 2	CAPM	0.248***	−0.205**	0.170**	−0.283***	0.255**	−0.424***	0.061	−0.078	0.200***	−0.055	0.071	−0.037
		(0.088)	(0.091)	(0.078)	(0.081)	(0.113)	(0.117)	(0.066)	(0.069)	(0.058)	(0.060)	(0.049)	(0.050)
	FF	0.026	0.148**	−0.006	0.012	−0.032	0.034	0.028	−0.004	0.070	0.135***	−0.029	0.112***
		(0.058)	(0.061)	(0.055)	(0.059)	(0.073)	(0.078)	(0.066)	(0.070)	(0.045)	(0.048)	(0.040)	(0.042)
	Carhart	0.030	0.167***	0.003	0.049	−0.029	0.048	0.040	0.051	0.066	0.118***	−0.029	0.111**
		(0.058)	(0.062)	(0.054)	(0.058)	(0.073)	(0.079)	(0.063)	(0.068)	(0.045)	(0.048)	(0.040)	(0.043)
Portfolio 3	CAPM	0.173**	−0.066	0.243***	−0.163**	0.314***	−0.329***	0.144***	0.126**	0.232***	−0.137*	0.043	−0.052
		(0.069)	(0.071)	(0.069)	(0.071)	(0.099)	(0.102)	(0.053)	(0.055)	(0.070)	(0.072)	(0.066)	(0.067)
	FF	0.016	0.184***	0.089*	0.072	0.074	0.041	0.076	0.228***	0.073	0.108*	−0.100**	0.172***
		(0.051)	(0.054)	(0.053)	(0.057)	(0.071)	(0.075)	(0.051)	(0.054)	(0.053)	(0.056)	(0.049)	(0.052)
	Carhart	0.016	0.181***	0.091*	0.082	0.073	0.035	0.073	0.216***	0.076	0.120**	−0.097*	0.184***
		(0.051)	(0.055)	(0.054)	(0.058)	(0.071)	(0.076)	(0.051)	(0.055)	(0.053)	(0.057)	(0.049)	(0.053)
Portfolio 4	CAPM	0.102*	0.011	0.304***	−0.013	0.393***	−0.251***	0.251***	0.038	0.210**	−0.145*	0.170*	−0.200**
		(0.061)	(0.063)	(0.061)	(0.063)	(0.091)	(0.095)	(0.063)	(0.066)	(0.083)	(0.085)	(0.090)	(0.092)
	FF	0.005	0.180***	0.180***	0.164***	0.175**	0.074	0.183***	0.118*	0.018	0.164**	−0.018	0.108
		(0.052)	(0.055)	(0.051)	(0.054)	(0.068)	(0.072)	(0.061)	(0.065)	(0.059)	(0.063)	(0.067)	(0.071)
	Carhart	0.008	0.191***	0.173***	0.134**	0.173**	0.069	0.170***	0.057	0.023	0.187***	−0.014	0.125*
		(0.052)	(0.056)	(0.050)	(0.053)	(0.068)	(0.073)	(0.058)	(0.062)	(0.059)	(0.063)	(0.067)	(0.072)
Portfolio 5	CAPM	0.024	0.087**	0.379***	0.068	0.541***	−0.042	0.288***	0.117	0.037	−0.364***	0.557***	−0.459***
		(0.034)	(0.035)	(0.067)	(0.069)	(0.093)	(0.097)	(0.072)	(0.074)	(0.109)	(0.112)	(0.127)	(0.130)
	FF	0.057*	0.046	0.278***	0.192***	0.346***	0.234***	0.225***	0.183**	−0.134*	−0.020	0.302***	−0.029
		(0.034)	(0.036)	(0.060)	(0.064)	(0.077)	(0.081)	(0.070)	(0.074)	(0.079)	(0.083)	(0.096)	(0.101)
	Carhart	0.059*	0.053	0.266***	0.138**	0.339***	0.202**	0.212***	0.125*	−0.121	0.038	0.307***	−0.010
		(0.034)	(0.036)	(0.057)	(0.062)	(0.076)	(0.082)	(0.067)	(0.072)	(0.076)	(0.082)	(0.096)	(0.102)

Note: *, **, and *** represent significance levels at 10%, 5%, and 1%, respectively.
SAI is the search attention index. The variable definitions are reported in Table S1 (available online). The standard errors are reported in the parentheses. CAPM is the capital asset, pricing model. FF stands for Fama and French Model. Carhart is the Carhart four-factor model.

and highly significant for stocks with a low value and positively significant for stocks with a high value. Further, the results strongly confirm that unexplained variations in the excess returns of the portfolios based on value are observed in the models.

In addition to the market, size, and value factors, the SAI also explains variations in the returns in India. The results of Table 4 show that when the SAI is low, the abnormal returns are low, and vice-versa. When the coefficients of the SAI increase from portfolio1 to portfolio5, the abnormal returns of the portfolios increase in the same direction. Further, the SAI is one of the factors that explains stock returns for growth stocks under the CAPM and value stocks under the Fama and French (1993) and Carhart (1997) models. The results support the price pressure hypothesis proposed by Barber and Odean (2008), which states that individual investors are net buyers of "attention-grabbing" stocks. The retail investors' search attention increases when they search for relevant financial information on the stocks for the purpose of investment. As a result, the returns are high in the short run but low in the long run. When retail investors buy attention-grabbing stocks, their price increases in the short run and decreases in the long run. The reason is that investors do not search for information when they sell stocks because they already own them and therefore know about them. Therefore, stocks that capture investor attention and for which investors search information intensively tend to generate excess returns and high trading volumes temporarily. This study confirms that value stocks capture investor attention in the market. Therefore, arbitrage opportunities may arise for both high-value and low-value stocks in the Indian stock market.

4.6. Size-value portfolios and double sorting with the search attention index (SAI)

First, we divide the sample stocks based on their market capitalization. Small-cap stocks are in the bottom 50 percent of the market capitalization, and large-cap stocks are in the top 50 percent of the market capitalization. Then, the two groups of stocks are divided into five quantiles, based on the price-to-book value and we form 10 portfolios. The primary purpose of this study is to examine whether the asset-pricing models with the SAI explain variations in stock returns. Table 4 (see Table S4, available online) shows that stocks with a small size and lowest value earn a significantly negative return, whereas stocks with a small size and the highest value have a significantly abnormal return. In the Fama and French (1993) and Carhart (1997) models, the coefficients are high for stocks with a small size and the highest value. This shows that the effect of investor attention is high for stocks that earn higher returns. In other words, retail investors search for stocks that yield higher returns. That is, when the SAI increases, the excess returns increase for stocks with a small size and the highest value in the short run and decrease in the long run. These results further support the price pressure hypothesis proposed by Barber and Odean (2008) and confirm the "size effect" and the investor attention effect, which explain variations in stock returns. Additionally, Table 4 (see Table S5, available online) shows that the alpha coefficients are higher for stocks with a large size and higher value. The portfolio of small stocks with a low value earns returns from −0.136 to −0.425, whereas the returns of the portfolio with a small size and higher value are from 0.346 to 0.541. In the group of large stocks, the alpha coefficients are significantly negative for stocks with a lower value and significantly positive for those with a higher value. This further confirms that the value effect is stronger for small stocks in the market (Fama and French, 2016) and that the SAI effect is higher for value stocks when the value effect exists.

4.7. Beta portfolios and the search attention index (SAI)

Baker et al. (2011), Schneider et al. (2020), and others show that stocks with a low beta have higher returns than those with a high beta. To investigate the beta effect, we divide the stocks into five quantiles based on their beta, from *beta1* for stocks with the lowest beta to *beta5* for those with the highest beta. We estimate the time-series returns for each portfolio from September 2012 to July 2017, consisting of 257 weekly observations. We then employ the asset-pricing models by incorporating the SAI as one of the explanatory variables. Table 4 (see Table S6, available online) presents the results for each portfolio sorted by the beta. When the beta is low, the market, size, and momentum factors explain the variations in the stock returns. At the same time, all three models fail to fully capture variations in the excess returns for stocks with an extremely low beta.

The SAI is negatively related to the excess returns of stocks for all portfolios in the CAPM model, whereas a significantly positive relationship is observed for the portfolios in the Fama and French (1993) and Carhart (1997) models. The results reveal that when the coefficients of the SAI from the portfolio1 to the portfolio4 increase, their alpha coefficients also increase. However, when the SAI is high, the excess returns are high up to the *beta4* portfolio. If the beta is high, the stock sensitivity with respect to the market is also high. Therefore, variability in the stock returns is high. Thus, retail investors in the Indian stock market who fear taking high risk will select stocks with a moderate beta.

4.8. Volatility portfolios and the search attention index (SAI)

Next, to test the volatility effect in the markets, we sort the sample stocks into five quantiles by the 12-week volatility of the returns. First, we calculate the standard deviation for each stock for the past 12 weeks. For every week, we divide the sample stocks into five quantiles from *portfolio1* to *portfolio5* based on the 12-week volatility, in which *portfolio1* is a portfolio containing stocks with low volatility and *portfolio5* contains stocks with high volatility. We then apply the asset-pricing models to capture the factors that explain variations in the excess returns of the volatile stocks. Table 4 (see Table S7, available online) reports the results of the models and reveals that the coefficients associated with the market, size, and value factors increase from portfolio1 to portfolio5, whereas the coefficients associated with the SAI increase up to portfolio4 and then decreases. Further, when the coefficients of the SAI are negative, the excess abnormal returns are positive, and vice-versa.

Table 4 also shows that the excess abnormal returns are high for stocks with high volatility and low for stocks with low volatility. Moreover, the coefficients on SAI increase until the portfolio4 and then decrease. This shows that even though returns are high for stocks with extremely high volatility, retail investors in the Indian stock market are risk averse. The coefficients associated with the SAI for portfolio2 and *portfolio3* are highly significant and indicate that investors prefer stocks with moderate risk levels. Baker and Wurgler (2007) and Joseph et al. (2011) argue that the investor sentiment index and abnormal returns increase from the bottom to the top portfolio of stocks sorted by volatility. Nevertheless, we find that investor attention is irrelevant for stocks with extremely high risk in the market. Moreover, we confirm that SAI is one of the factors that explains variations in the excess return of stocks.

4.9. Search attention index (SAI), abnormal returns and their reversal

Next, we investigate the explanatory power of the SAI in the long run. First, we form a portfolio that consists of a long position in the top quintile of stocks sorted by Google search intensity (Q5) and a short position in the lowest quintile of stocks sorted by Google search intensity (Q1). We then calculate the average return on the stocks every week, and the results are presented in Table 6. The first-week average return for a long-short portfolio is 0.384 percent, and the next four-week average return is 0.39 percent. The highest average

Table 6
Long-run returns from the long-short portfolio based on search intensity.

Week	Average return
Week 1	0.384
Weeks 2−4	0.390
Weeks 5−8	0.525
Weeks 9−12	0.181
Weeks 13−16	0.083

Table 5
Returns on long-short portfolios (N = 257).

	Models	Intercept	RM_RF	SMB	HML	WML	SAI	Adj R^2
Long on small stocks and short on large stocks (S1−S5)	CAPM	0.326***	1.546***				−0.664***	0.160
		(0.121)	(0.263)				(0.125)	
	FF	−0.017	0.475***	3.426***	1.533***		−0.135*	0.750
		(0.067)	(0.176)	(0.165)	(0.133)		(0.071)	
	Carhart	−0.019	0.485***	3.412***	1.560***	0.092	−0.144**	0.750
		(0.067)	(0.177)	(0.166)	(0.140)	(0.139)	(0.073)	
Long on high-value stocks and short on low-value stocks (V5−V1)	CAPM	0.585***	−3.607***				0.420***	0.394
		(0.128)	(0.278)				(0.132)	
	FF	0.674***	−1.627***	0.044	−2.592***		0.164	0.632
		(0.102)	(0.266)	(0.249)	(0.201)		(0.108)	
	Carhart	0.644***	−1.460***	−0.191	−2.139***	1.514***	0.030	0.708
		(0.091)	(0.238)	(0.224)	(0.188)	(0.186)	(0.098)	
Long on small & high-value and short on small & low-value stocks (SV5-SV1)	CAPM	0.677***	−2.824***				0.515***	0.327
		(0.117)	(0.254)				(0.121)	
	FF	0.771***	−1.200***	−0.215	−2.135***		0.277***	0.546
		(0.098)	(0.256)	(0.240)	(0.194)		(0.104)	
	Carhart	0.749***	−1.079***	−0.387*	−1.806***	1.101***	0.179*	0.598
		(0.092)	(0.242)	(0.227)	(0.191)	(0.190)	(0.099)	
Long on large & high-value and short on large & low-value stocks (BV5-BV1)	CAPM	0.590***	−3.805***				0.218	0.340
		(0.152)	(0.331)				(0.157)	
	FF	0.569***	−1.949***	1.257***	−2.383***		0.116	0.512
		(0.134)	(0.350)	(0.327)	(0.265)		(0.142)	
	Carhart	0.530***	−1.733***	0.952***	−1.796***	1.961***(0.247)	−0.059	0.608
		(0.120)	(0.315)	(0.296)	(0.248)		(0.129)	
Long on low-beta and short on high-beta stocks (beta1-beta5)	CAPM	0.167	−4.510***				0.326***	0.551
		(0.117)	(0.255)				(0.121)	
	FF	0.247**	−2.795***	−0.004	−2.246***		0.100	0.709
		(0.096)	(0.252)	(0.236)	(0.191)		(0.102)	
	Carhart	0.227**	−2.682***	−0.163	−1.939***	1.026***	0.009	0.739
		(0.091)	(0.240)	(0.225)	(0.189)	(0.188)	(0.098)	

Note: *, **, and *** represent significance levels at 10%, 5%, and 1%, respectively. SAI is the search attention index. RM_RF is the weekly excess market returns. SMB is the small minus big stocks. HML is the high minus low stocks. WML is the winning minus and losing stocks. The variable definitions are reported in Table S1 (available online). The standard errors are reported in the parentheses. CAPM is the capital asset, pricing model. FF stands for Fama and French Model. Carhart is the Carhart four-factor model.

return is observed from the fifth to the eighth weeks. This shows that when the search intensity is high, the average return is low in the short run but high in the long run. Further, we estimate the alpha coefficient for stocks with high search intensity based on Fama and Macbeth's approach. The results are presented in Table S8 (available online) for 16 weeks, showing that the alpha coefficient is negative and increases slowly in the long run. The alpha coefficient in the fourth week is 3.06 and later decreases. In the fourteenth week, the alpha coefficient is high, with a value of 3.92. This shows that when investors buy stocks, abnormal returns increase for a short period and then decrease in the long run. These results are consistent with the results of Joseph et al. (2011). Fig. 1 and 2 (see Figure S1, available online) also confirm that the returns increase in the short run and then decrease in the long run.

4.10. Trading strategy and robustness test

Finally, we form a trading strategy based on firm characteristics, such as market capitalization, price-to-book value, beta, volatility, and Google search intensity. Initially, we observe that stocks with high search intensity (Q5) earn higher returns than stocks with lower search intensity (Q1). Our investment strategy involves short-selling low-search-intensity stocks and buying high-search-intensity stocks, and the results related to this portfolio are presented in Table 3. The estimated alpha coefficients of the portfolio for the different models are highly significant with values of 0.442, 0.459, and 0.439, respectively. The search attention-based investment strategy has highly significantly positive returns. Further, the success of the search attention-based strategy is relatively robust to selection of the Google search attention measure for portfolio construction.

Next, we form an investment strategy that involves going long on stocks with the lowest market capitalization and short on stocks with the highest market capitalization. The results for this portfolio are presented in Table 5. The estimated alpha coefficient for the CAPM model is 0.326 percent and highly significant at the 1 percent level. At the same time, the beta associated with the SAI is negative. This further confirms that search attention is negatively related to the excess returns on low-cap stocks.

We then form different portfolios based on trading strategies that involve going long on high-value stocks and short on low-value stocks, long on small stocks with a high value and short on small stocks with a low value, long on big stocks with a high value and short on big stocks with a low value, long on stocks with a low beta and short on stocks with a high beta,

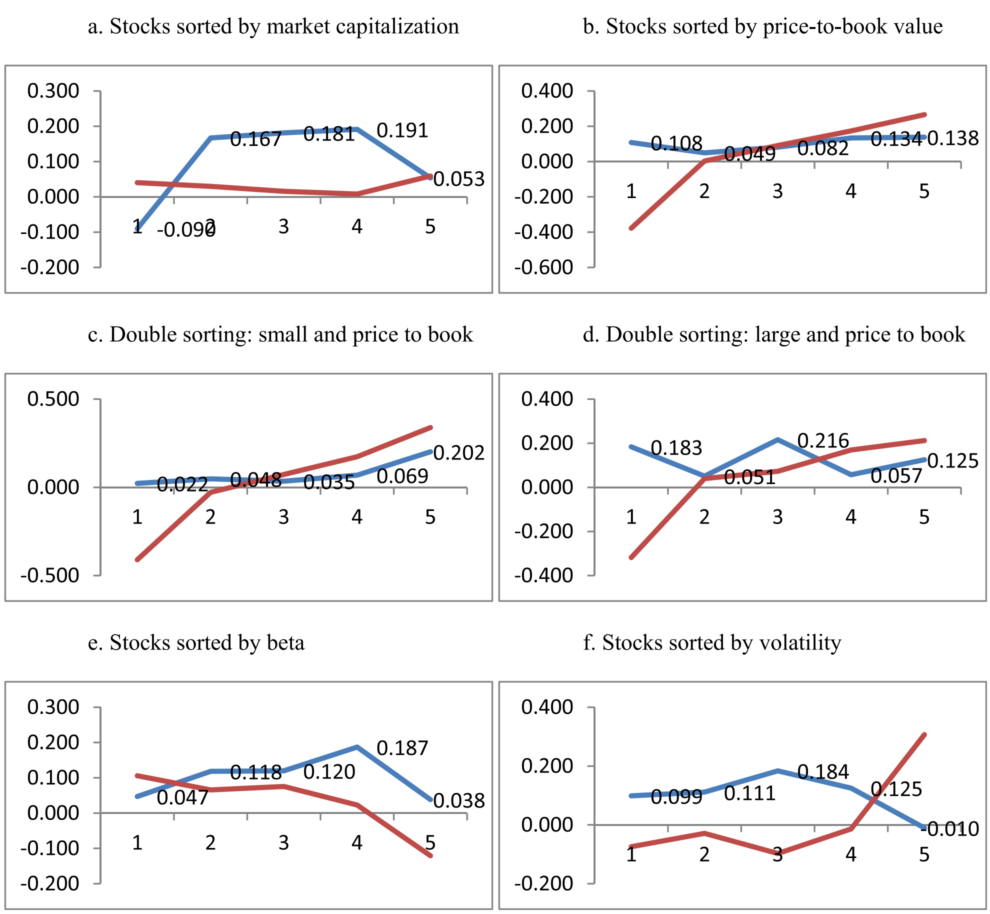

Fig. 1. Beta association with the SAI and firm characteristics. This figure explains the movement of the coefficient associated with the SAI and the alpha coefficient of the portfolios sorted by firm characteristics such as market capitalization, price-to-book value, beta, volatility, and search intensity over the period September 2012 to July 2017. The red line denotes the alpha coefficient and the blue line the beta coefficient associated with the SAI. a. Stocks sorted by market capitalization, b. Stocks sorted by price-to-book value.

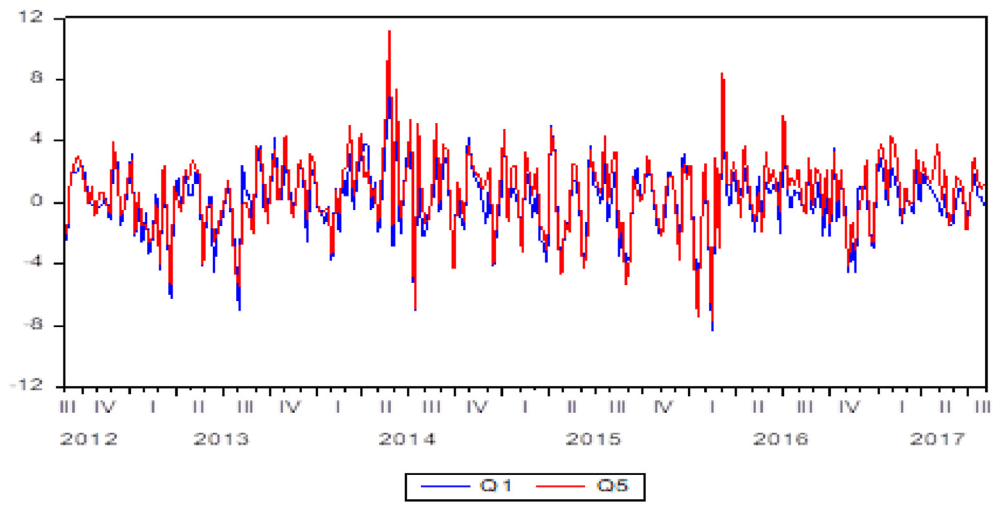

Fig. 2. Excess returns of the stocks with low search intensity (Q1) and stocks with high search intensity (Q5).

and long on high-volatility stocks and short on low-volatility stocks. The results of these long-short portfolios are also in Table 5. The results reveal that the estimated abnormal returns are highly significant and positively different from zero. Moreover, the success of this strategy is fairly robust to the type of portfolio. Further, the SAI is negatively related to low-cap stocks and positively related to stocks with a value and highly volatile stocks. Finally, we observe that in India, in addition to the market, size, value, and momentum factors, the SAI also explains variations in the excess returns of stocks.

5. Summary and conclusion

This study provides evidence-based results of the relationship between Google search intensity and stock returns in India. First, we introduce the SAI as a new factor and investigate whether it explains variations in stock returns when various asset-pricing models are used. Initially, we find that the raw average returns and estimated abnormal returns of the portfolios based on search intensity are significantly high for stocks with higher search intensity (Q5) and low for stocks with lower search intensity (Q1). We then find that excess returns are high when search intensity is low for small-cap stocks. Further, we find that the SAI profoundly affects the portfolio of stocks with a higher value, and vice-versa. This shows that the SAI positively influences the excess returns of value stocks. The results of double-sorting portfolios by size and value reveal that the effects of size, value, and investor attention explain variations in stock returns. This further confirms that the value effect is greater for small stocks (Fama and French, 2016). The results also reveal that the SAI effect is higher for value stocks when the value effect exists.

In addition, we report that when the coefficient associated with the SAI increases from low-to high-beta portfolios, abnormal returns increase in the same direction. Further, the results of the volatility portfolios show that the coefficients of the SAI increase until the portfolio4 of volatile stocks and then decrease. This shows that even though the return is high for stocks with extremely high volatility, retail investors in the Indian stock market are risk averse. Baker and Wurgler (2007) and Joseph et al. (2011) argue that investor sentiment and abnormal returns increase from the bottom to the top portfolio of stocks sorted by volatility. Nevertheless, we show that in India investor attention is irrelevant for stocks with extremely high risk.

Next, we find that when investors buy stocks, abnormal returns increase for a short period and then decrease in the long run. These results are consistent with those of Barber and Odean (2008) and Joseph et al. (2011), who show that when investors search for information on stocks, their buying behavior increases. As a result, the returns increase in the short run and then decrease in the long run. Finally, we design a trading strategy, whose results reveal that the estimated abnormal returns are highly significant and positively different from zero. Moreover, the success of this strategy is relatively robust to the type of portfolio. Further, the SAI is negatively related to small-cap stocks and positively related to stocks with a high value and high volatility. Ultimately, we observe that in India, in addition to the market, size, value, and momentum factors, the SAI also explains variations in excess stock returns.

Our findings offer constructive information about investor behavior, with a valuable measure of investor attention. Notably, our methodology can be used to form trading strategies that augment the risk management practices of listed companies. Applications of the SAI can provide better proxies for investor attention in other markets, such as gas, oil, energy, currency, and other commodities. Further, our conclusions lead us to suggest that investors restructure their portfolios by analyzing the patterns in Google search intensity.

We believe that our theoretical and methodological contributions enhance understanding of the characteristics and behavior of market returns from different perspectives. Despite the novelty of our study, it has limitations that offer directions for future research. One potential avenue of research is measurement of the attention effect at the crosscountry level, rather than in a single market. Our models could also be extended by creating an SAI based on other factors, novel variables, and samples from other regions.

Declaration of competing interest:

The authors declare no conflict of interest.

Acknowledgments

The authors are grateful to the editor and anonymous referees for insightful comments that significantly improved the paper. Further, we would like to thank Professor Vighneswara Swamy, IBS-Hyderabad, for his valuable comments on this research paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bir.2021.04.003.

References

Abedin, M. Z., Chi, G. T., Moula, F. E., Zhang, T., & Kabir Hassan, M. (2019). An optimized support vector machine intelligent technique using optimized feature selection methods: Evidence from Chinese credit approval data. *Journal of Risk Model Validation, 13*(2), 1–46.

Adachi, Y., Masuda, M., & Taketa, F. (2017). Google search intensity and its relationship to the returns and liquidity of Japanese startup stocks. *Pacific-Basin Finance Journal, 46*(B), 243–257. https://doi.org/10.1016/j.pacfin.2017.09.009

Afkhami, M., Cormack, L., & Ghoddusi, H. (2017). Google search keywords that best predict energy price volatility. *Energy Economics, 67*(9), 17–27. https://doi.org/10.1016/j.eneco.2017.07.014

Agarwalla, S. K., Jacob, J., & Varma, J. R. (2013). *Four factor model in Indian equities market. Working Paper No. 2013-09-05*. Indian Institute of Management Ahmedabad. https://faculty.iima.ac.in/~iffm/Indian-Fama-French-Momentum/four-factors-India-90s-onwards-IIM-WP-Version.pdf.

Baker, M., Bradley, B., & Wurgler, J. (2011). Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal, 67*, 40–54.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance, 61*(4), 1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives, 21*(2), 129–152.

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. *Financial Markets and Portfolio Management, 25*(3), 239–264. https://doi.org/10.1007/s11408-011-0165-y

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*(1), 3–18. https://doi.org/10.1016/0304-405X(81)90018-0

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies, 21*(2), 785–818. https://doi.org/10.1093/rfs/hhm079

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *Review of Financial Studies, 30*(9), 3009–3047. https://doi.org/10.1093/rfs/hhx031

Bijl, L., Kringhaug, G., Molnar, P., & Sandvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis, 45*(3), 150–156. https://doi.org/10.1016/j.irfa.2016.03.015

Carhart, M. M. (1997). On the persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57–82. https://doi.org/10.2307/2329556

Chan, L. K. C., Hamao, Y., & Lakonishok, J. (1991). Fundamentals and stock returns in Japan. *The Journal of Finance, 46*(5), 1739–1764. https://doi.org/10.2307/2328571

Chen, L., Lu, H., & Yang, L. (2013). Investor sentiment, disagreement, and the breath-return relationship. *Management Science, 59*(5), 1076–1091. https://doi.org/10.1287/mnsc.1120.1633

Chen, T. (2017). Investor attention and global stock returns. *Journal of Behavioral Finance, 18*(3), 358–372.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS: Investor sentiment and asset prices. *Review of Financial Studies, 28*(1), 1–32. https://doi.org/10.1093/rfs/hhu072

Da, Z., Engelberg, J., & Goa, P. (2011). In search of attention. *The Journal of Finance, 665*(5), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Das, P., & Ziobrowski, A. (2015). The relationship between Indian realty stocks and online searches. *Journal of Emerging Market Finance, 14*(1), 1–19. https://doi.org/10.1177/0972652714567994

Dowling, M., & Lucey, B. M. (2005). Weather, biorhythms, beliefs and stock returns: Some preliminary Irish evidence. *International Review of Financial Analysis, 14*(3), 337–355. https://doi.org/10.1016/j.irfa.2004.10.003

Edmans, A., Garcia, D., & Norli, O. (2007). Sports sentiment and stock returns. *The Journal of Finance, 62*(4), 1009–1032. https://doi.org/10.1111/j.1540-6261.2007.01262.x

Fama, E. (1970). Efficient capital market: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417. https://doi.org/10.2307/2325486

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427–465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics, 116*(1), 1–22.

Fama, E. F., & French, K. R. (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies, 29*(1), 69–103.

Gustavo, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breath of ownership, and liquidity. *Review of Financial Studies, 17*(2), 439–461. https://doi.org/10.1093/rfs/hhg039

Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance, 58*(3), 1009–1032. https://www.jstor.org/stable/3094570.

Jagadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from an online search. *International Journal of Forecasting, 27*(4), 1116–1127. https://doi.org/10.1016/j.ijforecast.2010.11.001

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kaplanski, G., & Levy, H. (2010). Sentiment and stock prices: The case of aviation disasters. *Journal of Financial Economics, 95*(2), 174–201. https://doi.org/10.1016/j.jfineco.2009.10.002

Kim, N., Lucivjansk, K., Molnar, P., & Villa, R. (2018). Google searches and stock market activity: Evidence from Norway. *Finance Research Letters, 28*, 208–220. https://doi.org/10.1016/j.frl.2018.05.003

Merton, R. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance, 42*(3), 483–510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x

Metwa, N., Kabir Hassan, M., & Elhoseny, M. (2017). Genetic algorithm-based model for optimizing bank lending decisions. *Expert Systems with Applications, 80*(1), 75–82.

Mnif, E., Salhi, B., Kabir Hassan, M., & Jarboui, A. (2020). Big data tools for Islamic financial analysis. *Intelligent Systems in Accounting, Finance and Management, 27*(1), 10–21.

Palomino, F., Renneboog, L., & Zhang, C. (2009). Information salience, investor sentiment, and stock returns: The case of British soccer betting. *Journal of Corporate Finance, 15*(3), 368−387. https://doi.org/10.1016/j.jcorpfin.2008.12.001

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics, 80*(3), 563−602.

Schneider, P., Wagner, C., & Zechner, J. (2020). Low-risk anomalies? *The Journal of Finance, 75*(5), 2673−2718.

Siganos, A. (2013). Google attention and target price run ups. *International Review of Financial Analysis, 29*(1), 219−226. https://doi.org/10.1016/j.irfa.2012.11.002

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics, 50*(3), 665−690.

Smith, G. P. (2012). Google internet search activity and volatility prediction in the market for foreign currency. *Finance Research Letters, 9*(2), 103−110. https://doi.org/10.1016/j.frl.2012.03.003

Stattman, D. (1980). Book values and stock returns. *The Chicago MBA: A Journal of Selected Papers, 4*, 25−45.

Swamy, V., & Dharani, M. (2019). Investor attention using the Google search volume index- impact on stock returns. *Review of Behavioral Finance, 11*(1), 55−69. https://doi.org/10.1108/RBF-04-2018-0033

Swamy, V., Dharani, M., & Taketa, F. (2019). Investor attention and Google search volume index: Evidence from an emerging market using quantile regression analysis. *Research in International Business and Finance, 50*(4), 1−17. https://doi.org/10.1016/j.ribaf.2019.04.010

Taketa, F., & Wakao, T. (2014). Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pacific-Basin Finance Journal, 27*(1), 1−18. https://doi.org/10.1016/j.pacfin.2014.01.003

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in Asia-Pacific stock markets. *Pacific-Basin Finance Journal, 38*, 107−124. https://doi.org/10.1016/j.pacfin.2016.03.008

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139−1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Venkataraman, M., Panchapagesan, V., & Jalan, E. (2018). Does internet search intensity predict house prices in emerging markets? A case of India. *Property Management, 36*(1), 103−118. https://doi.org/10.1108/PM-01-2017-0003

Vlastakis, N., & Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking & Finance, 36*(6), 1808−1821. https://doi.org/10.1016/j.jbankfin.2012.02.007

Ying, Q., Kong, D., & Luo, D. (2015). Investor attention, institutional ownership, and stock return: Empirical evidence from China. *Emerging Markets Finance and Trade, 51*(3), 672−685. https://doi.org/10.1080/1540496X.2015.1046339

Yuan, Y. (2008). *Attention and trading*. Financial Institutions Center, Wharton School, University of Pennsylvania.

Yung, K., & Nafar, N. (2017). Investor attention and the expected returns of REITs. *International Review of Economics & Finance, 48*(1), 423−439. https://doi.org/10.1016/j.iref.2016.12.009