Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Information dissemination across global markets during the spread of COVID-19 pandemic

Abhinava Tripathi, Ashish Pandey

ABSTRACT

This study examines the information dissemination process across 25 major global market indices during the times of COVID-19 pandemic spread. The results suggest that the information from non-systematic sources contributed to the price decline and increased volatility. In contrast, the systematic information lowered the volatility and facilitated the recovery process towards more stable markets. These results have important implications for policymakers and regulators in the development of efficient markets.

1. Introduction

Financial markets provide the conduit to channel the resources of an economy to the highest valued use, and in turn, promote growth and development (Chinn & Ito, 2006; Greenwood & Smith, 1997; Stiglitz, 1989). Significant efforts have been made by regulators, academics, investors, and other market participants to improve this ability of financial markets – i.e., market efficiency. The timely dissemination of reliable information on the firm-specific and macroeconomic events that affect the security fundamentals through future cash flows is essential to generate homogenous investor expectations, and consequently, higher levels of market efficiency. Over the last 40 years, there has been considerable progress in this regard (e.g., the introduction of electronic limit-order books, algorithmic trading). However, the prior evidence suggests that there have been intermittent events of crises with prolonged effects on global markets. These events include the great depression of 1930s, the oil crisis of 1970s, the recession of 2000, the sub-prime mortgage crisis of 2008, and the European sovereign debt crisis of 2009-12. The extant literature provides evidence that these crisis events are characterized by highly volatile financial markets and security prices that are far away from the efficient levels, over considerably long horizons (Aloui et al., 2011; Bai et al., 2012; Brutti, 2011; Calice et al., 2013; Campello et al., 2010; Taylor, 2009). The crisis periods involve uncertainty and increased information asymmetry that result in heterogeneous investor expectations. These heterogeneous investor expectations lower the levels of market efficiency and drive down the prices away from informationally efficient values. The prices are expected to remain subdued until the market participants absorb the full information content of the long term implications of the crisis. In this backdrop, the recent financial crisis driven by coronavirus pandemic (‘COVID-19’) is of particular interest.

* We would like to thank the Editor, Professor Brian Lucey, and the two anonymous reviewers for their constructive comments that substantially improved the paper. We are responsible for remaining errors.
* Corresponding author.
E-mail addresses: abhinava.tripathi@ms.iitr.ac.in (A. Tripathi), ashish.pandey@iiml.ac.in (A. Pandey).

https://doi.org/10.1016/j.iref.2021.02.004
Received 20 June 2020; Received in revised form 19 December 2020; Accepted 4 February 2021
Available online 9 February 2021
1059-0560/© 2021 Elsevier Inc. All rights reserved.
International Monetary Fund estimates that the economic recession from COVID-19 is likely to be worse than the 2008 financial crisis (IMF, 2020). In the absence of a reliable mechanism to either prevent the viral infection or cure the infection, governments worldwide have resorted to extreme measures such as countryside lockdowns. These measures have long-term implications on economic activity from supply as well as demand side. However, a preliminary examination of prices and volatility across global equity markets (reported in online Appendix A, B, and C) suggest that the equity prices have recovered and reached a relative calm, after witnessing a sharp decline in the four-week window starting from the third week of February 2020 to the third week of March 2020. The sharp recovery in the equity prices is contrasting to the patterns observed during earlier major crisis events and indicates that the market participants were able to understand and absorb the economic implications of this crisis to a significant extent in a relatively short time horizon. The theory of market microstructure proposes that the price movements are determined by the arrival of the new information and the process that disseminates this information into market prices (Andersen, 1996). In this paper, we examine the role of information sources in the price formation process during the COVID-19 crisis. Specifically, we investigate the effect of systematic and non-systematic components of information on return and volatility of 25 major global stock market indices.1

We posit that the COVID-19 crisis is unique in the sense that, across the world, its real impact (in terms of affliction, loss of life, and economic activity) is continuously monitored and evaluated in two ways that are distinct from the earlier crises. First, a number of headline indicators2 for COVID-19 and their analysis by the academic community in the popular media are available to the investors on a real-time basis. These headline indicators (HIs) are expected to cause substantial information dissemination and aggregation. The availability of information on a real-time basis enables investors to assess the expected future implications for economic activity. These headline indicators serve as a systematic source of common information and contribute to the formation of homogenous investor price expectations (Andersen, 1996; Chae, 2005; Figlewski, 1978, 1981). Second, investors have access to random non-systematic information from the internet (google search). This source of information, however, may be less reliable and accurate, and may contribute to the noise in prices. Kumar and Shah (2018) document the role of misinformation, disinformation, and opinion-based views that result in information inaccuracy. We argue that the novelty of the virus strain and the complex disease transmission mechanism may contribute to the distortion of information by a naive commentator on the internet, even without the intent to deceive. The high degree of technical prowess necessary to process the technical minutiae of the virus may result in misinformation. This misinformation is then spread unconsciously to others via articles, blogs, and social media posts. The non-systematic information from the internet, therefore, may contribute to uncertainty and heterogeneity in the expectations of market participants, leading to volatile markets. These two drivers behind the information flow may have a contrasting effect on the prices and volatility of financial markets. The systematic component of the information flow should attenuate the volatility and contribute to the recovery process by causing uniformity of expectations. In contrast, the non-systematic component of information flow should accentuate the volatility in the financial markets and result in informationally less efficient prices.

We use the daily global COVID-19 affliction and death rates as the HIs to capture the systematic component of information flow. We employ the google search volume index (“GSVI”) measure taken from google trends for different keywords related to the COVID-19 pandemic to capture the non-systematic component of information flow. Prior evidence suggests that the information demand represented by GSVI contributes to investor sentiment and leads to informationally less efficient prices (Da et al., 2011, 2014). We examine the contribution of systematic and non-systematic components to the information aggregation process across 25 global markets. The results are consistent with our initial hypotheses. The systematic component of the information available through HIs (total deaths and total cases) plays an important role in information dissemination and leads to informationally efficient prices. The systematic information lowers the volatility in markets and facilitates the recovery process. In contrast, the non-systematic component of information proxied using GSVI measure contributes to higher levels of volatility and sentiment-driven prices that are informationally less efficient. We also employ a novel Bayesian time-series approach to examine the robustness of our results (Scott, 2017; Scott & Varian, 2014, 2015). Lastly, using the quantile regression framework, we also document that the relationship between returns and investor sentiment is particularly significant across the lower quantile of returns (negative returns). This evidence is consistent with some of the previous studies (Badshah, 2013; Hibbert et al., 2008; Tantaopas et al., 2016), which suggest that the effect of investor sentiment on prices is considerably higher in falling markets. These results have important implications for academicians, policymakers, and regulators interested in containing and mitigating the repercussions of crises on financial markets.

The rest of the paper is structured as follows. Section 2 provides background and literature review. Sections 3 and 4 discuss the data, sample, variables, and methodology. The empirical analysis of results and robustness tests are provided in Section 5, and Section 6 concludes the study.

---

1 In the context of this study, the terms ‘systematic’ and ‘non-systematic’ reflect the nature of the source of information. For example, the information obtained through google search would comprise non-systematic information, as it would be highly idiosyncratic and dependent upon the individual user search queries. In contrast, the information like ‘total cases’, which can be obtained from reliable sources (like www.worldometers.info), is considered a systematic information. Because this information is expected to be identical even if obtained from multiple credible sources, a wider consensus is expected on this information as well as its implications.

2 For example, the website https://www.worldometers.info/coronavirus/#countries provides a number of Covid-19 related parameters on day to day basis. It is important to note that these headline indicators are made available by a number of credible organizations. All of these sources provide accurate, and therefore, largely identical data (such as global death and affliction). This data is continuously monitored and evaluated across the world resulting in similar expectations about the future global macroeconomic conditions.
2. Background and related literature

The fact that investor sentiment (or “animal spirits”) can affect prices and cause deviations from fundamental asset values has been long recognized (Keynes, 1936; Merton, 1987). In the last thirty years, however, the notion has gained considerable prominence with the emergence of an important stream of literature known as “Behavioral Finance” (See for example: Andrei & Hasler, 2014; Baker & Wurgler, 2006; Barber et al., 2008; Barber & Odean, 2007; Barberis et al., 1998; Bondt & Thaler, 1985; De Long et al., 1990; Stambaugh et al., 2014, 2012; Yu & Yuan, 2011). The majority of these studies examine the impact of investor sentiment on the rational asset pricing models, and endeavor to explain the anomalous behavior that contradicts the predictions of “Efficient Market Hypothesis” (EMH: Fama, 1970, 1991, 1998). For example, Yu and Yuan (2011) show that, during sentiment-driven markets, prices move away from informationally efficient values and weaken the conventional positive mean-variance relationship (i.e., risk-return trade-off). Literature suggests that markets are characterized by high sentiment and volatility during crisis periods. In these situations, arbitrage activity becomes difficult, and therefore, the effect of sentiment on prices can persist over a short-to medium-term (De Long et al., 1990). The sentiment-driven noise traders introduce noise into prices and drive prices away from informationally efficient values (Brown & Cliff, 2004).

With this background, we note that two particular aspects of investor behavior are widely documented: overreaction and underreaction to news (Barberis et al., 1998). First, the overreaction hypothesis is supported by the behavioral heuristic known as “representativeness” (Tversky & Kahneman, 1974). The representativeness heuristic suggests that individuals view recently observed events as a representation of the entire population of events and ignore the probabilistic aspects. Owing to this, investors overreact to the news considerably more than that would have been posited by EMH. Generally, this kind of behavior is observed during crisis situations characterized by uncertainty and volatile prices (Bondt & Thaler, 1985). In these situations, markets are dominated by uninformed and fearful noise traders who overreact and precipitate the fall in prices. These overreactions tend to get corrected, and a price reversal is witnessed over the short-to medium-term. Second, the underreaction hypothesis derives the theoretical underpinnings from the “conservatism” aspect of individual behavior (Edwards, 1968, 1982). The underreaction hypothesis suggests that individuals are risk-averse and are particularly sensitive to the risk of losing money. Therefore, they underreact to the news. Generally, this kind of behavior is observed during recovery periods when sentiments are low (Antoniou et al., 2013; Barber & Odean, 2000; Barberis et al., 1998; Daniel et al., 1998; Hong & Stein, 1999). In these situations, markets are dominated by rational risk-averse investors, as the sentiment-driven investors stay away from the market. These investors act cautiously to the news that contradicts their valuation, and thus underreact to the arrival of new information. The implication of the hypothesis is a gradual recovery process with autocorrelated price movements. In this study, we employ the hypotheses of overreaction and underreaction to explain the behavior of global markets during COVID-19.

More recently, Google has publicly made available the search volume index (SVI) data constructed from internet users’ google-search queries (Choi & Varian, 2009a, 2009b). Since then, several studies have employed this SVI as a proxy of retail investor attention (Bank

| Table 1 |
|---|
| Details of the market indices included in the study. |
| **Country** | **Index** | **Stock Exchange** |
| **American stock exchanges** | | |
| U.S.A. | NASDAQ Composite | NASDAQ |
| | S&P500 | NYSE, NASDAQ |
| | Dow Jones Industrial Average (DJIA) | NYSE, NASDAQ |
| | NYSE Composite | NYSE |
| Canada | S&P/TSX Composite | Toronto stock exchange |
| Brazil | IBovespa | Sao Paulo stock exchange |
| **European stock exchanges** | | |
| U.K. | Financial Times Stock Exchange (FTSE)-100 | London stock exchange |
| Ireland | ISEQ 20 | Euronext-Dublin |
| Germany | Deutscher Aktienindex (DAX) performance index | Frankfurt stock exchange |
| France | Cotation Assistée en Continu (CAC 40) | Euronext-Paris |
| Sweden | OMX Stockholm 30 (OMX30) | Stockholm stock exchange |
| Switzerland | Swiss Market Index (SMI) | SIX Swiss Exchange |
| Russia | MOEX | Moscow stock exchange |
| Spain | IBEX35 | Bolsa de Madrid |
| Belgium | BEL 20 | Euronext-Brussels |
| Austria | Austrian Traded Index (ATX) | Vienna Stock Exchange |
| Norway | OSEAX | Oslo Stock Exchange |
| Italy | FTSE MIB | Borsa Italiana Milan |
| **Asia-Pacific region stock exchanges** | | |
| China | SZSE Composite | Shenzhen Stock Exchange |
| | SSE Composite | Shanghai Stock Exchange |
| Japan | Nikkei-225 | Tokyo stock exchange |
| India | Nifty-500 | National stock exchange |
| Australia | BSE-SENSEX | Bombay stock exchange |
| South-Korea | ASX | Australian securities exchange |
| | Korea Composite Stock Price (KOSPI) | Korea Stock Exchange |
et al., 2011; Da et al., 2011, 2014; Ding & Hou, 2015; Joseph et al., 2011; Smith, 2012; Takeda & Wakao, 2014; Tantaopas et al., 2016; Vlastakis & Markellos, 2012; Vozlyublennaia, 2014). For example, Da et al. (2011, 2014) employ the google search volume index (GSVI) measure to proxy the sentiment of uninformed retail investors. Da et al. (2014) construct a measure of negative sentiment (“FEARS” index) using keywords such as “unemployment,” “recession,” and “bankruptcy.” They argue that this measure captures the behavior of sentiment-driven noise traders. During crisis events, the noise traders dominate the markets, and cause an uncertain and volatile environment characterized by falling prices. They also document that the falling prices reflect these sentiment-driven traders’ overreaction; and therefore, prices exhibit predictable reversals in the short-to-medium-term. Andrei and Hasler (2014), Aouadi (2013), Vlastakis and Markellos (2012) and Vozlyublennaia (2014) similarly show that the volatility of stock returns and risk premium increase with an increase in investor attention and uncertainty evidenced by a higher google search volume. Following the extant literature, we also attempt to capture the effect of investor sentiment on the price formation and volatility during the COVID-19 pandemic through the GSVI measure.

A growing strand of literature suggests that the impact of news coming from differing sources may have different implications for financial markets (Brenner et al., 2009; Fedyk, 2018; Jiao et al., 2020; Mamaysky, 2020). For example, Jiao et al. (2020) document that the news from traditional media is considered by individuals as genuine information and leads to subsequent volatility declines. In contrast, the news from social media creates more noise and leads to higher volatility levels. Similarly, Brenner et al. (2009) argue that information from reliable sources resolves uncertainty and disagreement among market participants, leading to a decline in volatility levels. During COVID-19 pandemic, headline indicators (HIs) represented reliable and systematic information (as compared to the random non-systematic Google search by individuals), and should induce uniform expectations about the future world view among the global market participants. Therefore, we examine the contribution of this systematic information component to the information aggregation process during the pandemic. Tantaopas et al. (2016) document that the response of the market to upturns and downturns is asymmetric. They suggest that markets are much more sensitive to negative news (Badshah, 2013; Hibbert et al., 2008; Tantaopas et al., 2016). We also examine this aspect in more detail.

3. Data, sample, and variables

Daily price and volume data of market index for 25 major global stock market indices (Table 1) are taken from the “Datastream-Eikon” database. These indices cover the largest global financial markets that are severely impacted by the COVID-19 pandemic. The period of analysis spans from January 1, 2020 to April 30, 2020. The log-returns \( r_t = \ln(p_t/p_{t-1}) \) of an index are computed based on its closing values. The European Centre for Disease Prevention and Control provides the daily details of headline indicators (HI) on COVID-19. The results reported in Table 2 – the stationarity of the individual country returns and other variables (physical indicators, GSVI measures) were also tested using standard tests (ADF and PP tests). The results suggest that all the variables are stationary.

4. Methodology

We examine the effect of HI and GSVI on the price formation process (return and volatility), using the following empirical specifications.

\[
\begin{align*}
\text{specifications (1 a & b)}
\end{align*}
\]

Here, \( r_{t,t} \) and \( h_{t,t}^{2} \) denote the return and conditional volatility for the period ‘t-1’ to ‘t’ for the country-index ‘c’. Specifications (1 a & b) examine the effect of information arrival on returns. The inclusion of lags of return in specification (1b) is to control for the serial
descriptive statistics for the measures employed in the study for 25 indices, over the study period (January 01, 2020 to April 30, 2020). These measures are returns, physical indicators (total cases and deaths), GSVI measure (using the keyword “Coronavirus”). ‘N’ denotes the number of observations. All the measures are in the log-difference form (similar to the return computation). Also, the test statistics for cross-sectionally augmented Im, Pesaran and Shin (IPS) unit-root test for panel models are shown (implemented with “cipstest” command in R package “plm”), indicating the stationarity of the data. For computation purposes, HIs and GSVIs are replicated for all the corresponding return observations.

correlation in returns due to the expected component of information. These specifications are similar in construction to those of Chordia et al. (2008), Chordia and Subrahmanyam (2004), and Hasbrouck (1988). Mixture of distribution hypothesis (MDH) postulates that the arrival of information should have a relationship with the return-volatility (Epps & Epps, 1976; Harris, 1987). In the spirit of MDH, specification (2) examines the effect of information arrival on conditional volatility. The models are estimated using pool and panel fixed-effects methods. Heteroscedasticity and autocorrelation consistent robust standard errors are employed in the computation of t-statistics. Following the BIC criteria, we consider four lags (p = 4, q = 4) of the return, GSVI, and HI measures.

5. Empirical results

The time-series of price-volume (Candle-stick chart), returns, and conditional volatility [AR(4)-EGARCH (1,1) model] for each country are provided in the online Appendices A, B, and C, respectively. We observe similar dynamics across all markets in the sample set. First, the effect of the onset of the pandemic resulting in a decline in prices and an increase in conditional volatility is noticed from the third week of February 2020 till the third week of March 2020. The subsequent period witnessed a steady rise in prices along with a decrease in conditional volatility, suggesting the start of the recovery process. The evidence is consistent across all the markets in the sample. It is important to note that the extent of pandemic spread and fatality varies widely across countries (e.g., extremely high in the U.S. and the European markets and moderate to low in Asian markets). However, the price dynamics across markets is almost identical, demonstrating that in the European and the U.S. markets and moderate to low in Asian markets. However, the price dynamics across markets is almost identical, as if the prices are affected by a common information arrival process. The results from the model estimation based on specifications (1a & b) are presented in Tables 3 and 4. Table 3 reports the results using pooled regression approach, while Table 4 reports the results using panel fixed-effect methods. The results suggest a positive relationship between the HIs and returns, indicating that HIs contribute to the rise in prices (positive returns) and facilitate the recovery process. In contrast, the GSVI indicator exhibits a negative relationship with returns and contributes to the fall in prices. We also confirm the robustness of these results by estimating individual index specific regressions. The results, not reported for brevity, are qualitatively similar (sign of the coefficients) for all the countries.

Table 5 (Panels A and B) reports the results pertaining to the impact of systematic and non-systematic information components on conditional volatility, using specification (2). The results, using pooled regression and panel regression methods, highlight the linkages between the information arrival and volatility. The contemporaneous term of the variable representing the systematic component of the information exhibits a positive relationship with the conditional volatility, indicating the arrival of information. However, as this information is absorbed, markets become less volatile and more efficient, evidenced by the negative coefficient of the lagged terms representing the systematic component. In contrast, the results suggest that the non-systematic component of the information contributes to an increase in volatility. The systematic component augments market efficiency by facilitating homogenous expectations and lower volatility. In contrast, the non-systematic component of information leads to heterogenous expectations and higher volatility levels.

\[ \text{Test-statistic} = -2.95^{***} \]
\[ \text{Test-statistic} = -3.27^{***} \]
\[ \text{Test-statistic} = -5.70^{***} \]
\[ \text{Test-statistic} = -2.73^{***} \]

The results are robust to one, two, three, four, and higher lagged returns.

The implementation of the GARCH model in ‘R’ is offered by ‘rugarch’ package. We compared Standard-GARCH (1,1), EGARCH (1,1), and GJR-GARCH (1,1) models. The information criteria (Akaike, Bayes, Shibata, and Hann-Quinn) consistently indicate that EGARCH model offers the best-fit. These results, not shown here for brevity, are available on request from the authors.

Chinese markets exhibit similar behavior, but two weeks earlier than global markets.

10 For example, during the period employed in this study, the extent of pandemic was considerably less in Asian markets (like India) as compared to that in the European and the U.S. markets. However, the global price dynamics would suggest that markets have anticipated the impact of pandemic on Asian markets in the coming months.
Table 3
Pooled regressions of returns on headline indicator (HI) and GSVI.

Panel A: without controls Equation (1a)

|       | N   | HI_C | HI_lag1 | HI_lag2 | HI_lag3 | HI_lag4 | GSVI_C | GSVI_lag1 | GSVI_lag2 | GSVI_lag3 | GSVI_lag4 | Adj. R² | F-value   |
|-------|-----|------|---------|---------|---------|---------|--------|-----------|-----------|-----------|-----------|---------|-----------|
| TC    | 2,067 | 0.039*** | 0.009*** | 0.018*** | 0.006*** | -0.003 | -0.050*** | 0.002 | -0.011** | -0.018*** | -0.017*** | 9.93%   | 389.854*** |
|       |      | 7.571 | 2.897   | 8.873   | 2.781   | -1.071 | -10.981  | 0.322 | -2.373   | -8.606    | -7.025    |         |           |
| TD    | 2,067 | 0.03*** | 0.009*** | 0.026*** | 0.000   | 0.014*** | -0.06*** | -0.004 | -0.015*** | -0.017*** | -0.015*** | 11.76%  | 243.915*** |
|       |      | 8.893 | 6.056   | 13.321  | -0.140  | 7.352   | -11.100  | -0.833 | -3.006   | -7.330    | -5.367    |         |           |

Panel B: with controls Equation (1b)

|       | N   | HI_C | HI_lag1 | HI_lag2 | HI_lag3 | HI_lag4 | GSVI_C | GSVI_lag1 | GSVI_lag2 | GSVI_lag3 | GSVI_lag4 | Adj. R² | F-value   |
|-------|-----|------|---------|---------|---------|---------|--------|-----------|-----------|-----------|-----------|---------|-----------|
| TC    | 2,067 | 0.038*** | 0.014*** | 0.016*** | 0.007*** | -0.004 | -0.047*** | -0.012*** | -0.006 | -0.016*** | -0.019*** | 16.83%  | 412.525*** |
|       |      | 7.447 | 4.596   | 8.201   | 4.009   | -1.604 | -10.857  | -3.351 | -1.528   | -10.166   | -7.728    |         |           |
| TD    | 2,067 | 0.032*** | 0.016*** | 0.025*** | 0.004*** | 0.009*** | -0.056*** | -0.022*** | -0.011*** | -0.016*** | -0.016*** | 18.84%  | 275.132*** |
|       |      | 8.571 | 8.035   | 14.824  | 2.640   | 4.148  | -10.321  | -5.380 | -2.719   | -8.481    | -5.156    |         |           |

Note: (1) ‘N’ denotes the number of observations. (2) *, **, and *** denote statistical significance at 10%, 5%, and 1% levels. (3) The GSVI (“Coronavirus”) is selected as it exhibits the highest interest over time. Results with other GSVIs (“Corona,” “Covid-19,” and “Covid”) are qualitatively similar and not reported for brevity.

Panels A and B present the results from the pooled method [following Equations (1 a & b)]. We report the coefficients of headline indicator [HI: Total cases (TC)/Total deaths (TD)] and GSVI measure [Keyword (“Coronavirus”)], t-statistics (below), adjusted-R², and F-values. The coefficients are reported for the contemporaneous (C), and four lag terms (lag1, lag2, lag3, and lag4). Information criteria are employed for the selection of lags.
Table 4
Panel Fixed-effect regressions of returns on headline indicator (HI) and GSVI.

Panel A: without controls Equation (1a)

|       | TC     |   |   |   |   |   |   |   |   |   |   | Adj. $R^2$ | F-value     |
|-------|--------|---|---|---|---|---|---|---|---|---|---|-----------|-------------|
| N     | 2,067  |   |   |   |   |   |   |   |   |   |   |           |             |
| HI_C  | 0.039***| 0.01***| 0.018***| 0.006***| -0.003| -0.05***| 0.002| -0.011**| -0.018***| -0.017***| 8.92%      | 381.023***  |
| HI_lag1| 0.003   |   |   |   |   |   |   |   |   |   |   |           |             |
| HI_lag2| 2.916   | 8.858 | 2.931| -0.950| -11.017| 0.305| -2.378| -8.611 | -7.072     | 8.982      | 6.118      | 13.538     | -0.010      | 7.583       | -11.163     | -0.845       | -3.012       | -7.403       | -5.446     | 10.78%      | 238.192*** |
| HI_lag3|         |   |   |   |   |   |   |   |   |   |   |           |             |
| HI_lag4|         |   |   |   |   |   |   |   |   |   |   |           |             |
| GSVI_C |         |   |   |   |   |   |   |   |   |   |   |           |             |
| GSVI_lag1| 0.047***| -0.020***| 15.92%| 379.289***| 7.544 | 4.659 | 4.219 | -1.372 | -10.881 | -3.426 | -1.575 | -10.207 | 7.871     |
| GSVI_lag2| 0.004***| 0.009***| 0.026***| 0.000  | 0.015***| -0.06***| -0.004| -0.015***| -0.017***| -0.015***| 10.78%      | 238.192*** |
| GSVI_lag3| 0.056***| 0.012***| 0.023***| 0.011***| 0.017***| 0.016***| -0.020***| -0.017***| -0.017***| -0.017***| 17.98%      | 241.393*** |
| GSVI_lag4| 0.011***| 0.017***| 0.016***| 0.011***| 0.017***| 0.017***| 0.017***| 0.017***| 0.017***| 0.017***| 17.98%      | 241.393*** |

Panel B: with controls Equation (1b)

|       | TC     |   |   |   |   |   |   |   |   |   |   | Adj. $R^2$ | F-value     |
|-------|--------|---|---|---|---|---|---|---|---|---|---|-----------|-------------|
| N     | 2,067  |   |   |   |   |   |   |   |   |   |   |           |             |
| HI_C  | 0.039***| 0.014***| 0.017***| 0.007***| -0.003| -0.047***| -0.012***| -0.006| -0.016***| -0.020***| 15.92%      | 379.289***  |
| HI_lag1| 7.544   | 4.659 | 8.211| 4.219 | -1.372| -10.881| -3.426| -1.575| -10.207 | 7.871     |
| HI_lag2| 0.004***| 0.009***| 0.026***| 0.000  | 0.015***| -0.06***| -0.004| -0.015***| -0.017***| -0.015***| 10.78%      | 238.192*** |
| HI_lag3| 0.056***| 0.012***| 0.023***| 0.011***| 0.017***| 0.016***| -0.020***| -0.017***| -0.017***| -0.017***| 17.98%      | 241.393*** |
| HI_lag4| 0.011***| 0.017***| 0.016***| 0.011***| 0.017***| 0.017***| 0.017***| 0.017***| 0.017***| 0.017***| 17.98%      | 241.393*** |
| GSVI_C |         |   |   |   |   |   |   |   |   |   |   |           |             |
| GSVI_lag1|         |   |   |   |   |   |   |   |   |   |   |           |             |
| GSVI_lag2|         |   |   |   |   |   |   |   |   |   |   |           |             |
| GSVI_lag3|         |   |   |   |   |   |   |   |   |   |   |           |             |
| GSVI_lag4|         |   |   |   |   |   |   |   |   |   |   |           |             |

Panels A and B present the results from the panel fixed-effect method [following Equations (1 a & b)]. We report the coefficients of headline indicator [HI: Total cases(TC)/Total deaths (TD)] and GSVI measure [Keyword (“Coronavirus”)], t-statistics (below), adjusted-$R^2$, and F-values. The coefficients are reported for the contemporaneous (C), and four lag terms (lag1, lag2, lag3, and lag4). Information criteria are employed for the selection of lags.

Note: (1) ‘N’ denotes the number of observations. (2) *, **, and *** denote statistical significance at 10%, 5%, and 1% levels. (3) The GSVI (“Coronavirus”) is selected as it exhibits the highest interest over time. Results with other GSVIs (“Corona,” “Covid19,” and “Covid”) are qualitatively similar and not reported for brevity.
Table 5
Regressions of conditional volatility on headline indicator (HI) and GSVI.

| Panel A: Pooled model | N  | HI_C  | HI_lag1 | HI_lag2  | HI_lag3 | HI_lag4 | GSVI_C | GSVI_lag1 | GSVI_lag2 | GSVI_lag3 | GSVI_lag4 | Adj. R² | F-value |
|-----------------------|----|-------|---------|----------|---------|---------|--------|-----------|-----------|-----------|-----------|--------|---------|
| TC                    | 2,067 | 0.014*** | 0.000  | −0.008*** | 0.002 | −0.004** | 0.001*** | 0.006*** | 0.003*** | −0.010*** | 0.000 | 41.89% | 81.147*** |
|                       | 8,069 | 0.062 | −0.014 | 1.030  | −2.343 | 5.067  | 9.585 | 8.250 | −8.198 | 0.299 |
| TD                    | 2,067 | 0.028*** | 0.014*** | −0.001*** | −0.022*** | −0.008*** | −0.001*** | 0.001 | 0.004*** | 0.002* | −0.003*** | 60.37% | 117.353*** |
|                       | 8,524 | 0.715 | −7.416 | −8.413 | −5.548 | −5.736 | 1.420 | 5.814 | 1.904 | −6.429 |

| Panel B: Panel fixed-effects model | N  | HI_C  | HI_lag1 | HI_lag2  | HI_lag3 | HI_lag4 | GSVI_C | GSVI_lag1 | GSVI_lag2 | GSVI_lag3 | GSVI_lag4 | Adj. R² | F-value |
|-----------------------------------|----|-------|---------|----------|---------|---------|--------|-----------|-----------|-----------|-----------|--------|---------|
| TC                                | 2,067 | 0.014*** | 0.000  | −0.008*** | 0.001 | −0.003* | 0.001*** | 0.006*** | 0.003*** | −0.009*** | 0.000 | 41.23% | 31.203*** |
|                                   | 8,039 | 0.710 | −4.607 | 0.714  | −1.787 | 4.936  | 9.928 | 8.925 | −8.967 | −0.316 |
| TD                                | 2,067 | 0.027*** | 0.013*** | 0.001*** | −0.022*** | −0.008*** | −0.001*** | 0.001* | 0.004*** | 0.002* | −0.004*** | 63.46% | 71.833*** |
|                                   | 8,811 | 0.052 | −8.364 | −8.436 | −6.112 | −4.996 | 1.855 | 5.994 | 1.836 | −6.243 |

Panels A and B present the results from pooled and panel fixed-effect methods [following Equation (2)]. The conditional volatility is obtained using AR(4)-EGARCH (1,1) model on the time-series of returns. The model is selected using information criteria (AIC/BIC/SIC). The results from other GARCH models (standard GARCH and GJR-GARCH) are qualitatively similar and not reported for brevity. We report the coefficients of headline indicator [HI: Total cases (TC)/Total deaths (TD)] and GSVI measure [Keyword (“Coronavirus”)], t-statistics (below), adjusted-R², and F-values. The coefficients are reported for the contemporaneous (C), and four lag terms (lag1, lag2, lag3, and lag4). Information criteria are employed for the selection of lags.

Note: (1) ‘N’ denotes number of observations. (2) *, **, and *** denote statistical significance at 10%, 5%, and 1% levels. (3) The GSVI (“Coronavirus”) is selected as it exhibits the highest interest over time. Results with other GSVIs (“Corona”, “Covid19”, and “Covid”) are qualitatively similar and not report for brevity.
similar to those obtained from the pooled and panel regression models. The inclusion probabilities shown in Fig. 1 demonstrate the state variables evolves over time, following the process specified in equation (4). The error terms \( \epsilon_t \) and \( \eta_t \) are independent and Gaussian. The structural parameters of the model are denoted by arrays \( Z_T \), \( T_t \), and \( Q_t \). The state is defined by the three components, namely, a trend \( (\mu_t) \), a seasonal pattern \( (\delta_t) \), and a contemporaneous regression component \( (\beta'X_t) \), google trend measure (GSVI), and headline indicators (HI). The trend includes the long-run slope \( (\delta') \) along with a short-run autoregressive component defined by the parameter \( \rho \). The model adapted to this context becomes a “local level model”.

\[
\begin{align*}
    r_t &= \mu_t + \tau_t + \beta'X_t + \epsilon_t \\
    \alpha_{t+1} &= T_t \alpha_t + Q_t \eta_t
\end{align*}
\]

(3)

(4)

The observed data \( r_t \) (the index return series) is linked to a vector of latent state variables ‘\( \alpha_t \)’ through equation (3). This vector of state variables evolves over time, following the process specified in equation (4). The error terms \( \epsilon_t \) and \( \eta_t \) above are independent and Gaussian. The structural parameters of the model are denoted by arrays \( Z_T \), \( T_t \), and \( Q_t \). The state is defined by the three components, namely, a trend \( (\mu_t) \), a seasonal pattern \( (\delta_t) \), and a contemporaneous regression component \( (\beta'X_t) \), google trend measure (GSVI), and headline indicators (HI). The trend includes the long-run slope \( (\delta') \) along with a short-run autoregressive component defined by the parameter \( \rho \). The model adapted to this context becomes a “local level model”.

\[
\begin{align*}
    r_t &= \mu_t + \tau_t + \beta'X_t + \epsilon_t \\
    \mu_{t+1} &= \mu_t + \delta_t + \eta_t \\
    \delta_{t+1} &= D + \rho(\delta_t - D) + \eta_t \\
    \tau_{t+1} &= -\sum_{s=1}^{S-1} \tau_t + \eta_t
\end{align*}
\]

(5)

(6)

(7)

(8)

5.1. Bayesian structural time-series (BSTS) approach

To test the robustness of our results for the impact of systematic and non-systematic components of information on returns, we use the Bayesian structural time-series (BSTS) approach (Scott, 2017; Scott & Varian, 2014, 2015). The BSTS approach offers four key advantages as compared to traditional time-series techniques. First, the approach offers the flexibility to choose among the relevant state components, including trend, seasonality, and regression effects (Scott & Varian, 2014). Second, the Kalman filter used in the model gathers statistics about the time series as it moves forward through the list of elements, while the Kalman smoother moves backward through time and distributes information about later observations to earlier elements (Durbin & Koopman, 2012). These “general message passing algorithms” make the estimation of the model more informationally efficient. The third advantage offered by the BSTS model is the ability to use spike and slab priors. By employing spike and slab priors, a model with a large set of correlated variables is transformed into a parsimonious model that incorporates prior beliefs. Lastly, the application of the Bayesian model averaging method in the BSTS approach combines the feature selection results and prediction calculation (Hoeting et al., 1999). Using the Bayesian averaging technique, one can incorporate parameter uncertainty as well as model uncertainty through the prior distribution. The resulting model produces more reliable intervals for the forecasts. Based on their analysis of simulated and real stock market data, Jammalamadaka et al. (2018) document that the BSTS model offers better prediction accuracy when compared to the autoregressive integrated moving average regression (ARIMAX) model and the multivariate ARIMAX (MARIMAX) model. The structural model is defined by the following set of equations.

\[
r_t = Z_T^T \alpha_t + \epsilon_t
\]

\[
\alpha_{t+1} = T_t \alpha_t + Q_t \eta_t
\]

The observed data \( r_t \) (the index return series) is linked to a vector of latent state variables ‘\( \alpha_t \)’ through equation (3). This vector of state variables evolves over time, following the process specified in equation (4). The error terms \( \epsilon_t \) and \( \eta_t \) above are independent and Gaussian. The structural parameters of the model are denoted by arrays \( Z_T \), \( T_t \), and \( Q_t \). The state is defined by the three components, namely, a trend \( (\mu_t) \), a seasonal pattern \( (\delta_t) \), and a contemporaneous regression component \( (\beta'X_t) \), google trend measure (GSVI), and headline indicators (HI). The trend includes the long-run slope \( (\delta') \) along with a short-run autoregressive component defined by the parameter \( \rho \). The model adapted to this context becomes a “local level model”.

\[
r_t = \mu_t + \tau_t + \beta'X_t + \epsilon_t
\]

\[
\mu_{t+1} = \mu_t + \delta_t + \eta_t
\]

\[
\delta_{t+1} = D + \rho(\delta_t - D) + \eta_t
\]

\[
\tau_{t+1} = -\sum_{s=1}^{S-1} \tau_t + \eta_t
\]

The observed data \( r_t \) (the index return series) is linked to a vector of latent state variables ‘\( \alpha_t \)’ through equation (3). This vector of state variables evolves over time, following the process specified in equation (4). The error terms \( \epsilon_t \) and \( \eta_t \) above are independent and Gaussian. The structural parameters of the model are denoted by arrays \( Z_T \), \( T_t \), and \( Q_t \). The state is defined by the three components, namely, a trend \( (\mu_t) \), a seasonal pattern \( (\delta_t) \), and a contemporaneous regression component \( (\beta'X_t) \), google trend measure (GSVI), and headline indicators (HI). The trend includes the long-run slope \( (\delta') \) along with a short-run autoregressive component defined by the parameter \( \rho \). The model adapted to this context becomes a “local level model”.

\[
r_t = \mu_t + \tau_t + \beta'X_t + \epsilon_t
\]

\[
\mu_{t+1} = \mu_t + \delta_t + \eta_t
\]

\[
\delta_{t+1} = D + \rho(\delta_t - D) + \eta_t
\]

\[
\tau_{t+1} = -\sum_{s=1}^{S-1} \tau_t + \eta_t
\]

The observed data \( r_t \) (the index return series) is linked to a vector of latent state variables ‘\( \alpha_t \)’ through equation (3). This vector of state variables evolves over time, following the process specified in equation (4). The error terms \( \epsilon_t \) and \( \eta_t \) above are independent and Gaussian. The structural parameters of the model are denoted by arrays \( Z_T \), \( T_t \), and \( Q_t \). The state is defined by the three components, namely, a trend \( (\mu_t) \), a seasonal pattern \( (\delta_t) \), and a contemporaneous regression component \( (\beta'X_t) \), google trend measure (GSVI), and headline indicators (HI). The trend includes the long-run slope \( (\delta') \) along with a short-run autoregressive component defined by the parameter \( \rho \). The model adapted to this context becomes a “local level model”.

\[
r_t = \mu_t + \tau_t + \beta'X_t + \epsilon_t
\]

\[
\mu_{t+1} = \mu_t + \delta_t + \eta_t
\]

\[
\delta_{t+1} = D + \rho(\delta_t - D) + \eta_t
\]

\[
\tau_{t+1} = -\sum_{s=1}^{S-1} \tau_t + \eta_t
\]

The observed data \( r_t \) (the index return series) is linked to a vector of latent state variables ‘\( \alpha_t \)’ through equation (3). This vector of state variables evolves over time, following the process specified in equation (4). The error terms \( \epsilon_t \) and \( \eta_t \) above are independent and Gaussian. The structural parameters of the model are denoted by arrays \( Z_T \), \( T_t \), and \( Q_t \). The state is defined by the three components, namely, a trend \( (\mu_t) \), a seasonal pattern \( (\delta_t) \), and a contemporaneous regression component \( (\beta'X_t) \), google trend measure (GSVI), and headline indicators (HI). The trend includes the long-run slope \( (\delta') \) along with a short-run autoregressive component defined by the parameter \( \rho \). The model adapted to this context becomes a “local level model”.

\[
r_t = \mu_t + \tau_t + \beta'X_t + \epsilon_t
\]

\[
\mu_{t+1} = \mu_t + \delta_t + \eta_t
\]

\[
\delta_{t+1} = D + \rho(\delta_t - D) + \eta_t
\]

\[
\tau_{t+1} = -\sum_{s=1}^{S-1} \tau_t + \eta_t
\]
5.2. Asymmetric relationship between investor sentiment and returns

We estimate model 1(a & b) for the positive and negative returns separately. The corresponding results with the panel fixed-effects method are shown in Appendix D (Tables D1 and D2). The results are qualitatively similar to those reported in Tables 3 and 4. More interestingly, the explanatory power (Adjusted-$R^2$) of the model is considerably high during market downturns (negative returns) as compared to market upturns (positive returns). This suggests that the negative sentiment has a stronger relationship with trading behavior as compared to that of the positive sentiment. Some of the recent studies have documented a similar asymmetric relationship between returns and investor sentiment (Tantaopas et al., 2016).

In addition, we also examine this relationship in a quantile regression framework (Koenker, 2013; Koenker & Bassett, 1978). The quantile regression approach is particularly suitable for modeling the tail behavior of returns characterized by heteroscedasticity, skewness, and excess kurtosis. Moreover, the approach can appropriately handle nonlinearity in the relationship that is expected to

---

11 We thank the anonymous reviewer for suggesting these additional tests
12 Results for pooled regressions are qualitatively similar and not reported for brevity.
emerge during crisis situations such as the COVID-19 pandemic. The results corresponding to equation (1a) are shown in Appendix E. The results from equation (1b) are consistent and not reported for brevity. Overall, the results remain qualitatively similar to those reported in Tables 3 and 4, and Appendix D. These results also confirm the asymmetric nature of the relationship between returns and investor sentiment as indicated by the goodness-of-fit measure (Koenker & Machado, 1999).

Overall, our results appear to be the manifestation of the following dynamics between the noise and informed traders (Black, 1986). The noise traders are usually small investors (e.g., retail traders) with limited resources and information processing capabilities. These investors follow simple trading heuristics (like buy-low sell-high) and are affected by the attention-grabbing unsystematic news similar to that available through google search (Andrei & Hasler, 2014; Barber & Odean, 2007). In contrast to the noise traders, the informed traders are large institutional investors. They invest considerable time and resources in acquiring valuable information (Grossman & Stiglitz, 1980). During the crisis periods, fear sentiment is high, and markets are characterized by a dominant presence of noise traders. Informed traders prefer to stay away from these uncertain and volatile markets. The noise traders drive-down the prices to significantly lower levels, as compared to the informationally efficient levels (overreaction hypothesis). Given the dominance of these sentiment-driven noise traders in crisis events, their contribution to prices - higher volatility and falling prices - is captured by the relation of GSVI measure with returns and conditional volatility. Subsequently, as the prices witness large declines, the noise traders are forced to leave the market. At this point, the informed traders enter the market to carry out arbitrage activity, as they find the prices to be attractive and significantly away from their informationally efficient levels. As more and more informed traders enter the market, markets become less volatile, and prices recover to become more informationally efficient. This is captured by the relation of HI measure with returns and conditional volatility (Fleming et al., 2006; Mougoue & Aggarwal, 2011). The reported results also confirm the asymmetric nature of the relationship between returns and investor sentiment.

6. Conclusion

In this paper, we examined the effect of systematic and non-systematic components of information on the return and volatility of 25 major global stock market indices. Our results show that the systematic component of the information leads to lower volatility and informationally efficient prices. In contrast, the non-systematic component of information contributes to higher levels of volatility and sentiment-driven prices that are informationally less efficient.

The results presented in this study offer important implications to academicians, practitioners, regulators, and policymakers. These results suggest that the speed and quality of information dissemination significantly contribute to market efficiency, and in turn, the recovery of financial markets from crisis events such as COVID-19. Market regulators and policymakers may considerably improve the levels of market efficiency and facilitate investor welfare by creating reliable platforms for information dissemination and aggregation that are accessible to market participants at a low cost. Furthermore, during times of crisis, regulators may nudge the volatile markets through appropriate signals that are more reliable. The results show that GSVI and HI have explanatory power on stock returns and conditional volatility. This suggests that investor attention may indeed be a priced variable (Da et al., 2011, 2014). Therefore, asset pricing models (e.g., Fama French 3-factor model) can be further augmented with measures of investor attention. Lastly, given the nature of the research design, the study covers a relatively short time span of 4 months. This remains an important limitation of the study, and therefore, the findings of the study are to be interpreted accordingly.

Funding acknowledgment

The authors received no financial support for the research, authorship, and publication of this article.

Acknowledgements

All persons who have made substantial contributions to the work reported in the manuscript (e.g., technical help, writing and editing assistance, general support), but who do not meet the criteria for authorship, are named in the Acknowledgements and have given us their written permission to be named. If we have not included an Acknowledgements, then that indicates that we have not received substantial contributions from non-authors.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.iref.2021.02.004.

References

Aloui, R., Alissa, M. S. B., & Nguyen, D. K. (2011). Global financial crisis, extreme interdependences, and contagion effects: The role of economic structure? Journal of Banking & Finance, 35(1), 130–141.
Anderson, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. The Journal of Finance, 51(1), 169–204.
Andrei, D., & Hasler, M. (2014). Investor attention and stock market volatility. Review of Financial Studies, 28(1), 33–72.
Antoniou, C., Doukas, J. A., & Subrahmanyam, A. (2013). Cognitive dissonance, sentiment, and momentum. Journal of Financial and Quantitative Analysis, 48(1), 245–275.
Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. Journal of Financial Economics, 104(2), 288–302.
Stambaugh, R. F., Yu, J., & Yuan, Y. (2014). The long of it: Odds that investor sentiment spuriously predicts anomaly returns. Journal of Financial Economics, 114(3), 613–619.
Stiglitz, J. E. (1989). Financial markets and development (JSTOR) Oxford Review of Economic Policy, 5(4), 55–68.
Takeda, F., & Wakao, T. (2014). Google search intensity and its relationship with returns and trading volume of Japanese stocks. Pacific-Basin Finance Journal, 27, 1–18.
Tantaopas, P., Padungaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in Asia-Pacific stock markets. Pacific-Basin Finance Journal, 38, 107–124.
Taylor, J. B. (2009). The financial crisis and the policy responses: An empirical analysis of what went wrong. National Bureau of Economic Research.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124–1131.
Vlastakis, N., & Markellos, R. N. (2012). Information demand and stock market volatility. Journal of Banking & Finance, 36(6), 1808–1821.
Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. Journal of Banking & Finance, 41, 17–35.
Yu, J., & Yuan, Y. (2011). Investor sentiment and the mean–variance relation. Journal of Financial Economics, 100(2), 367–381.
Aouadi, A, et al. (2013). Investor attention and stock market activity: Evidence from France. Economic Modelling, 35, 674–681.