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.
Impacts of COVID-19 local spread and Google search trend on the US stock market

Asim K. Dey\textsuperscript{a,b}, G.M. Toufiqul Hoque\textsuperscript{c}, Kumer P. Das\textsuperscript{d,}\textsuperscript{*}, Irina Panovska\textsuperscript{e}

\textsuperscript{a} Department of Mathematical Sciences, University of Texas at El Paso, El Paso, TX 79968, USA
\textsuperscript{b} Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA
\textsuperscript{c} Department of Mathematics, Lamar University, Beaumont, TX 77715, USA
\textsuperscript{d} The Office of Vice President for Research, Innovation, and Economic Development, University of Louisiana at Lafayette, Lafayette, LA 70504, USA
\textsuperscript{e} School of Economic, Political, and Policy Sciences, University of Texas at Dallas, Richardson, TX 75080, USA

\section*{A R T I C L E   I N F O}
Article history:
Received 31 December 2020
Received in revised form 14 July 2021
Available online 30 September 2021

Keywords:
Covid-19
Stock market
Temporal network
Abnormal price
Volatility
Causality

\section*{A B S T R A C T}
We develop a novel temporal complex network approach to quantify the US county level spread dynamics of COVID-19. We use both conventional econometric and Machine Learning (ML) models that incorporate the local spread dynamics, COVID-19 cases and death, and Google search activities to assess if incorporating information about local spread improves the predictive accuracy of models for the US stock market. The results suggest that COVID-19 cases and deaths, its local spread, and Google searches have impacts on abnormal stock prices between January 2020 to May 2020. Furthermore, incorporating information about local spread significantly improves the performance of forecasting models of the abnormal stock prices at longer forecasting horizons.

© 2021 Published by Elsevier B.V.

\section*{1. Introduction}
After the COVID-19 pandemic started spreading worldwide, the US stock market collapsed significantly with the S&P 500 dropping 38\% between February 24, 2020 and March 20, 2020.\textsuperscript{1} Historically, natural disasters and large geopolitical events have had negative effects on stock markets. For example, Cagle \cite{1}, Cavallo and Noy \cite{2}, Worthington and Valadkhani \cite{3}, Worthington \cite{4}, and Shan and Gong \cite{5} study the impact of natural disasters, e.g., hurricanes and earthquakes, on the stock market. Hudson and Urquhart \cite{6}, Schneider and Troeger \cite{7}, Chau et al. \cite{8}, Beaulieu et al. \cite{9}, and Huynh and Burggraf \cite{10} evaluate the effect of political uncertainty and war on the stock market. The influences of the outbreak of infectious diseases, e.g., Ebola and SARS, on the stock indices are assessed in \cite{11,12,13,14}. However, due to the unprecedented public health and financial circumstances, both assessing the effects of COVID-19 on the stock market and assessing which variables are useful for predicting movements in the stock market has been a challenge that has spurred many developments in the literature. A number of recent studies attempt to assess the impact of the COVID-19 outbreak on the stock market and on the economy in general. Nicola et al. \cite{15} provide a review on the socioeconomic effects of COVID-19 on individual aspects of the world economy, and highlight that the pandemic affected all aspects of society from increases in domestic violence and changes in the agricultural supply chain to a large negative effect on financial markets. They note that similar declines have occurred in other stock indices and in financial markets

\textsuperscript{*} Corresponding author.
E-mail address: kumar.das@louisiana.edu (K.P. Das).
\textsuperscript{1} Author’s calculation using data for the S&P index between February and March 2020.

https://doi.org/10.1016/j.physa.2021.126423
0378-4371/© 2021 Published by Elsevier B.V.
globally. Using data from national new sources and data for stock market jumps, Baker et al. [16] analyze the reasons why the U.S. stock market reacted so much more adversely to COVID-19 than to previous pandemics that occurred in 1918–1919, 1957–1958 and 1968. The shift towards services and the large effects of the pandemic on the service sector led to large negative effects on both the real sector and the financial sector.

Onali [17], Zaremba et al. [18], Arias-Calluari et al. [19], and Cao et al. [20] perform statistical modeling to analyze the effect of COVID-19 on the stock market price and volatility. Arias-Calluari et al. [19] propose a number of methods to forecast movements in financial markets during the COVID-19 period. Onali [17] uses national data for cases and deaths and GARCH and vector autoregressive models for seven large economies, and finds mixed evidence in favor of negative effects of COVID-19 cases on the US stock market, apart from the number of reported cases for China. However, there is evidence of a positive impact, for some countries, on the conditional heteroscedasticity of the Dow Jones and S&P 500 returns, indicating a potential link between COVID-19 cases and stock market volatility. Similarly, Cao et al. [20] use national level data for cases and a panel of 14 stock indices and calculate the elasticity of stock indices to COVID-19 cases. They find that stock markets have a negative response to local (country-level) cumulative cases, and they move with local and non-local spreads, indicating a need for international efforts of containment to pare market losses.

Wagner [21] notes the link between stock market volatility and rise of COVID-19 globally, and highlights the need for flexible policies and investment strategies in the future. Zaremba et al. [18] use a panel regression to study how government interventions affected stock market volatility in 67 different countries, and find that non-pharmaceutical interventions such as event cancellations significantly increased market volatility. Shehzad et al. [22] use a nonlinear autoregressive distributed lag model and national level data to study the effect on stock markets. In contrast to Zaremba et al. [18], they find that deficiencies in the health sector can also directly cause negative effects on the stock market that go beyond the direct effects of the associated economic crisis. Rubaniy et al. [23] use national level data for European markets, find that lockdowns nor the stringent measures taken by the governments to improve effect of COVID-19 on European stock markets were effective. However, some of the financial measures by central banks (e.g., reduction in capital buffers) help mitigating the adverse impacts of COVID-19 on European stock markets. In contrast, Cox et al. [24] use a formal structural economic model to study the effects of policy announcements on the stock market during the early period of the pandemic and find that movements in the stock market were more reflective of sentiment than of substance.

The effects of the pandemic on financial markets go beyond the direct initial drop in the stock market. Bouri et al. [25] use a time-varying-parameter vector autoregression for multiple asset classes, and find that connectedness between financial assets increased drastically during the COVID-19 outbreak, raising concerns about potential financial contagion. This period was also connected with large spikes in uncertainty. Bouri and Gupta [26] construct multiple measures of uncertainty based on newspaper and on internet searches, and show that Bitcoin is a hedge against uncertainty for alternative measures of uncertainty. Dey and Das [27] evaluate the effects of mobility restrictions and social distancing on the crude oil price. Similarly, Gupta et al. [28] show that US treasury securities can hedge the risks associated with the financial market in the wake of the current COVID-19 pandemic.

Given the large initial decline in the stock market, and the close and increased connection between different classes of assets during this period, quantifying the effects of COVID-19 on the stock market and assessing what COVID-19 variables are important for predicting movements in the stock market is important both for investors and for policy makers. Despite the plethora of recent research studies that focus on this topic, there are still a number of important questions that need to be investigated. For example,

1. How can we quantify the local spread dynamics of COVID-19 and do the local spread, e.g., US county level spread of COVID-19, affect stock prices?
2. Do the number of COVID-19 cases and deaths influence stock prices?
3. Do the local spread, the number of COVID-19 cases and deaths influence the stock market’s volatility?
4. Do the Google search volumes related to COVID-19 exhibit any relationship with the stock price and volatility?
5. Do the local spread of COVID-19, the number of cases and deaths, and Google search volumes convey any additional information about the stock price dynamics, given more conventional economic variables?

The economic forecasting literature has shown that the inclusion of additional state and country levels predictors that capture regional and local variables substantially improves nowcasts and forecasts for aggregate economic activity (GDP and employment) Hernández-Murillo and Owyang [29], with the gains typically being concentrated during periods of large negative movements Owyang et al. [30]. Similarly, Dalton [31] finds that local spreads, as measured by the local incidence of the virus, has large effects on economic activity.

However, most of the previous literature that has studied the effects of COVID-19 on stock market prices and volatility has focused on the effects of national or global variables. Because local spread may be indicative both of disruption in local economic activity and might be linked to local sentiment, we augment the standard aggregate models with variables that capture the local spread, and we assess to what extent local spread affects the stock market price and volatility.

In order to quantify the local spread of COVID-19, we introduce the concept of temporal network and network motifs. This allows us to utilize multi-source information to quantify the impact of the spread on the US stock market. The
network structure allows us to leverage much richer data set that includes information not only about the total number of cases and cases within a county, but also about the spread across counties over time.\(^2\)

In addition to including controls for movements in previous prices and volatility and variables for COVID-19 spread, we also control for investor sentiment. Investor sentiment is another crucial determinant of stock market dynamics and can serve as an important policy transmission channel. Cox et al. [24] use a dynamic asset pricing model and high-frequency policy announcement news to study the effects of policy on the stock market during the COVID-19 crisis, and find that movements during the crisis have been more reflective of sentiment than substance, with the response to sentiment being a more important driver in the stock market than the responses to the actual monetary policy actions. This has also been highlighted in the fiscal literature, going back to Ramey [32], who points out that including policy measures that were anticipated as predictors in a time-series model can lead to biased estimates for the effects of fiscal policy. She suggests using narrative measures, forecasted values or forecast errors, or sentiment at time \( t \) as predictors rather than using the measure of actions that were implemented at time \( t \) that were expected in previous periods.

However, quantifying investor sentiment is not an easy task because of its unobservable and heterogeneous behaviors García Petit et al. [33], Gao et al. [34], Baker and Wurgler [35], Bandopadhaya and Jones [36]. In recent years, due to data availability, Google search volume has become a popular index of investor sentiment Bijl et al. [37], Kim et al. [38], Preis et al. [39]. Bollen et al. [40] determine Twitter feeds as the moods of investors and use the Twitter mood to predict the stock market. Alanyali et al. [41] Schumaker and Chen [42], Bomfim [43], and Albuquerque and Vega [44] evaluate the relationship between financial news and the stock market and find that news related to the asset significantly impact the corresponding stock price and volatility. In this study we use Google Trends as a proxy for investor sentiment and we also include Google Trends for monetary and fiscal policy actions to capture the policy anticipation.

In line with the previous literature, we find that Google search activities have contemporaneous relationship and predictive power on abnormal stock prices and volatility. Much more importantly, we leverage the use of temporal networks and network motifs to also include a measure of local COVID-19 spreads. We find that local spreads have a contemporaneous relationship with stock prices and that incorporating the local spreads leads to significant improvements when it comes to modeling abnormal stock prices relative to models that only incorporate national-level variables.

The rest of the paper is organized as follows. Section 2 describes the data, constructs a temporal network for COVID-19 spread, and defines the variables used in the study. The methodology is described in Section 3. Section 4 presents findings and a discussion of the results. Finally, Section 5 provides a conclusion.

2. Data and variables

The S&P 500 closing price from June 3, 2019 to May 29, 2020 data are obtained from Yahoo! Finance [45]. Google search data from January 2, 2020 to May 29, 2020 are obtained from Google Trends. We get US County level COVID-19 case data from the New York Times [46] and US county information from the US National Weather Service [47].

2.1. Abnormal stock price and volatility

We evaluate the impact of COVID-19 on abnormalities in the S&P 500 index. We define the daily abnormal S&P 500 price (\( AP_t \)) between January 2, 2020 and May 29, 2020 by subtracting the average price of the last seven months from the daily price and by dividing the resultant difference from the standard deviation of the last seven months (i.e., 148 days) as follows:

\[
AP_t = \frac{P_t - \frac{1}{148} \sum_{i=1}^{148} P_{t-i}}{\sigma_P},
\]

where, \( P_t \) is the daily closing price for day \( t \), \( \sigma_P \) is the standard deviation of the closing price in the last 148 days [37, 38]. We use daily squared log returns of prices \( P_t \) as a proxy for daily volatility (\( \text{Vol}_t \)) [48, 49]:

\[
r_t = \log \left( \frac{P_t}{P_{t-1}} \right), \quad \text{Vol}_t = r_t^2.
\]

2.2. COVID-19 cases

We study the impact of a number of COVID-19 variables (\( C \)), e.g., daily US total cases, daily US new cases, daily World total cases, etc. on \( AP_t \) and \( \text{Vol}_t \). For a complete list of COVID-19 variables see Table 1. We standardized each COVID-19 variable on the basis of a rolling average of the past 7 days and corresponding standard deviation as:

\[
CV_t = \frac{C_t - \mu_C}{\sigma_C},
\]

where, \( C_t \) is a COVID-19 variable (e.g., US total cases) at day \( t \), \( \mu_C \) and \( \sigma_C \) are the mean and standard deviation of the corresponding variable within the sliding window of days \([t - k, t - 1]\).

\(^2\) In this study we apply the concept of temporal networks and network motifs to model the local spread of COVID. However, as discussed in Section 3, these concepts can be used in many other empirical applications. Outside of disease dynamic modeling, our approach can also be applied to modeling financial or debt contagion.
2.3. Local spread through complex network analysis

A complex network represents a collection of elements and their inter-relationship. A network consists of a pair \( G = (V, E) \) of sets, where \( V \) is a set of nodes, and \( E \subseteq V \times V \) is a set of edges, \((i, j) \in E\) represents an edge (relationship) from node \( i \) to node \( j \). Here \(|V|\) is the number of nodes and \(|E|\) is the number of edges. The degree \( d_u \) of a node \( u \) is the number of edges incident to \( u \) i.e., for \( u, v \in V \) and \( e \in E \), \( d_u = \sum_{u \neq v} e_{u,v} \). A graph \( G' = (V', E') \) is a subgraph of \( G \), if \( V' \subseteq V \) and \( E' \subseteq E \). The largest connected component \((GC)\) is the maximal connected subgraph of \( G \). The elements of the \( n \times n \)-symmetric adjacency matrix, \( A \), of \( G \) can be written as

\[
A_{ij} = \begin{cases} 
1, & \text{if } (i, j) \in E \\
0, & \text{otherwise.} 
\end{cases}
\]  

A higher-order network structure, e.g., motif, represents local interaction pattern of the network. In a disease transmission a network motif provides significant insights about the spread of the diseases. For example, the presence of a dense motif or fully connected motif can increase the spread of the disease through the network, while a chain-like motif can decrease the spread of the disease [50]. A motif is a recurrent multi-node subgraph pattern. A detailed description of network motifs and their functionality in a complex network can be found in [51–54], and [55]. Fig. 1 shows all connected 3-node motifs \((T)\) and 4-node motifs \((M)\).

Temporal Network is an emerging extension of network analysis which appears in many domains of knowledge, including epidemiology [56–58], and finance [59–61]. A temporal network is a network structure that changes in time. That is, a temporal network can be represented with a time indexed graph \( G_t = (V(t), E(t)) \), where, \( V(t) \) is the set of nodes in the network at time \( t \), \( E(t) \subseteq V(t) \times V(t) \) is a set of edges in the network at time \( t \). Here \( t \) is either discrete or continuous. Fig. 2 depicts a small 15-node temporal network with time \( t = 1, 2, \) and \( 3 \).

In order to quantify the county level spread of COVID-19 we construct a complex network \((G_t)\) in each day \( t \) between Jan 2, 2020 to May 29, 2020: \( G = \{G_1, \ldots, G_r\} \), where \( r = 130 \). We evaluate the occurrences of different motifs in each \( G_t \). An increase number of motifs, i.e., \( T \) and \( M \), and other network features e.g., \( E \), indicate a higher spread in a local community. These increases of higher order network structures have potential impacts on \( AP \) and \( Vol \).

Let \( C \) be the set of counties in US, \( I \) is the set of COVID-19 new cases identified in \( C \) on a day \( t \), and \( D \) is the pairwise distance matrix in miles among centroid of the counties in \( C \). We use the following three steps to construct the COVID-19 spread network \((G_t)\) at time \( t \) and compute the occurrences of motifs in \( G_t \):

1. Each County in \( C \) with \( \gamma \) or more COVID-19 new cases, \( \gamma \in \mathbb{Z}^+ \), makes a node in the network \((G_t)\).
2. Two counties (i.e., nodes), \( i \) and \( j \), are connected by an edge if (1) both counties have \( \lambda \) or more COVID-19 new cases, \( \lambda \in \mathbb{Z}^+ \), and (2) the distance between \( i \) and \( j \) is less than \( \delta \), \( \delta \in \mathbb{R}_{>0} \). Therefore, the adjacency matrix, \( A^t \), is

![Fig. 1. All 3-node and 4-node connected network motifs.](image1)

![Fig. 2. A changing network shown over three time steps.](image2)
written as
\[
A_{ij}^t = \begin{cases} 
1, & \text{if } I_i^t, I_j^t > \lambda \& D_{ij} < \delta \\
0, & \text{otherwise}
\end{cases}
\] (5)

3. We compute occurrences of nodes \(V_t\), edges \(E_t\), different 3-node motif \(T(t)\), different 4-node motifs \(M(t)\), and size of the largest connected component \(GC(t)\) in \(G_t\).

   In this study, we choose \(\gamma = 5, \lambda = 5, \) and \(\delta = 100\). That is, if two counties both have 5 or more COVID-19 cases and if the distance between these two counties is less than 100 miles they are connected by an edge. However, any appropriate choice of the parameters \(\gamma, \lambda, \) and \(\delta\) can also be used to construct the COVID-19 spread network \((G_t)\). For illustration, Fig. 3 shows the COVID-19 spread network in US counties on April 11, 2020. We consider different network features e.g., \(E, T, M, \) etc. as metrics of the local spread of COVID-19. We normalize each of the network variables based on Eq. (3) as

\[
SP_t = \frac{N_t - \mu_N}{\sigma_N},
\] (6)

where, \(N_t\) is a network variable (e.g., \(E\)) at day \(t\), \(\mu_N\) and \(\sigma_N\) are the mean and standard deviation of the corresponding variable within the sliding window of days \([t - k, t - 1]\).

2.4. Google trend data

A number of studies, e.g., Bijl et al. [37], Preis et al. [62], and Kim et al. [38], show that there is a significant correlation between stock variables (e.g., return, volume, and volatility) and related Google searches, and Google search data can be used to predict future stock prices.

We investigate whether Google trend data affect the abnormal price, \(AP\), and volatility, \(Vol\). We obtain the volume of the COVID-19 related daily Google searches (e.g., “Coronavirus”) from Jan 2, 2020 to May 29, 2020. We select the location of a query in “US” and in the “World”. We standardized each Google search variable similar to Eq. (3) as

\[
GT_t = \frac{G_t - \mu_G}{\sigma_G},
\] (7)

where, \(G_t\) is a Google search variable at day \(t\), \(\mu_G\) and \(\sigma_G\) are the mean and standard deviation of the corresponding variable within the sliding window of days \([t - k, t - 1]\). Table 1 provides an overview of the data sets and variables that are used in this study.\(^3\)

3. Methodology

We investigate the impact of COVID-19 cases and deaths, local spread of COVID-19, and COVID-19 related Google search volumes on the abnormal stock price and volatility.

\(^3\) In a robustness check experiment, we also consider Google Trend searches for various economic policy variables, and the news based economic policy uncertainty index from [63] as additional explanatory variables. However, while these variables were significantly correlated with abnormal returns, they did not have a causal relationship at any lag. For brevity, the results for the additional policy variables are reported in Appendix.
3.1. Correlation and causality

A correlation test is widely used as an initial step to evaluate the relationship between the stock market and a potential covariate [38,39,41,62]. In this study, we use Spearman’s rank correlation to study correlation between stock data (AP and Vol) and each of the COVID-19 related variables.

To assess potential predictive utilities of COVID-19 cases, local spread, and Google search interests on abnormal price formation (AP) and Vol, we apply the concept of Granger causality [64]. Nonlinear Granger causality tests [65–67] are another alternative. However, such tests require substantially higher records of data. For example, Diks and Panchenko [66] show that in samples smaller than 500 observations their test may under-reject. Therefore because of having smaller records of data we use the classical linear Granger causality test.

The Granger causality test evaluates whether one time series is useful in forecasting another. It is very important to note that this test uses the statistical notion of causality. If variable X Granger-causes Y, this does not necessarily indicate a structural relationship in the microeconometric sense, but it does indicates that X has additional predictive power for Y. The term “causality” throughout the text refers to this statistical causality. Our focus in this study is to build a predictive model, and we assess if each variable is useful for predicting the financial variables of interest.

Let \( Y_t \), \( t \in Z^+ \) be a \( p \times 1 \)-random vector (\( AP_t \) or \( V_t \)) and let \( \mathcal{F}_Y^{(t)} = \sigma(\{Y_s : s = 0, 1, \ldots, t\}) \) denote a \( \sigma \)-algebra generated from all observations of \( Y \) in the market up to time \( t \). Consider a sequence of random vectors \( \{Y_t, X_t\} \), where \( X \) can be either COVID-19 cases, local spread or Google search volumes. Suppose that for all \( h \in Z^+ \)

\[
F_{t+h}(Y_t | \mathcal{F}_Y^{(t)}) = F_{t+h}(Y_t | \mathcal{F}_X^{(t-h)}),
\]

where \( F_{t+h}(Y_t | \mathcal{F}_Y^{(t)}) \) and \( F_{t+h}(Y_t | \mathcal{F}_X^{(t-h)}) \) are conditional distributions of \( Y_{t+h} \), given \( Y_{t-1}, X_{t-1} \) and \( Y_{t-1} \), respectively. Then, \( X_{t-1} \) is said not to Granger cause \( Y_{t+h} \) with respect to \( Y_{t-1} \). Otherwise, \( X \) is said to Granger cause \( Y \), which can be denoted by \( G_{X-Y} \), where \( \rightarrow \) represents the direction of causality [68–70].

We fit two models where one model includes \( X \) and the other does not include \( X \) (base model), and compare their predictive performance to assess causality of \( X \) to \( Y \) using an \( F \)-test, under the null hypothesis of no explanatory power in \( X \). For univariate cases we compare the following two models:

\[
y_t = \alpha_0 + \sum_{k=1}^{d} \alpha_k y_{t-k} + \sum_{k=1}^{d} \beta_k x_{t-k} + \epsilon_t,
\]

versus the base model

\[
y_t = \alpha_0 + \sum_{k=1}^{d} \alpha_k y_{t-k} + \tilde{\epsilon}_t.
\]

If \( \text{Var}(\epsilon_t) \) is significantly lower than \( \text{Var}(\tilde{\epsilon}_t) \), then \( X \) contains additional information that can improve forecasting of \( y \), i.e., \( G_{X-Y} \). We can also fit two linear vector autoregressive (VAR) models, with and without \( X \), respectively, and evaluate the statistical significance of model coefficients associated with \( X \).

3.2. Predictive models

To quantify the forecasting utility of the covariates (\( X \)), i.e., COVID-19 cases, US county level spread of COVID-19, and Google searches, we develop predictive models with and without \( X \) and compare their predictive performances. In order to conduct such a comparison, Box–Jenkins (BJ) class of parametric linear models are commonly used. However, different studies, e.g., Dey et al. [68] and Kane et al. [71], show that flexible Random Forest (RF) models often tend to outperform...
Table 2
Model description for abnormal price AP and varying predictors.

| Model  | Predictors |
|--------|-------------|
| Model $P_0$ | AP lag 1, AP lag 2, AP lag 3 |
| Model $P_1$ | AP lag 1, AP lag 2, AP lag 3, US total deaths lag 1, US total deaths lag 2, US total deaths lag 3, World new deaths lag 1, World new deaths lag 2, World new deaths lag 3 |
| Model $P_2$ | AP lag 1, AP lag 2, AP lag 3, Edges lag 1, Edges lag 2, Edges lag 3, GC lag 1, GC lag 2, GC lag 3, $T_1$ lag 1, $T_1$ lag 2, $T_1$ lag 3, $M_4$ lag 1, $M_4$ lag 2, $M_4$ lag 3 |
| Model $P_3$ | AP lag 1, AP lag 2, AP lag 3, “Covid-19” US lag 1, “Covid-19” US lag 2, “Covid-19” US lag 3, “Covid-19” US lag 4, “Covid-19” World lag 1, “Covid-19” World lag 2 |
| Model $P_4$ | AP lag 1, AP lag 2, AP lag 3, “Covid-19” US lag 1, “Covid-19” US lag 2, “Covid-19” US lag 3, “Covid-19” US lag 4, US total deaths lag 1, US total deaths lag 2 |

the BJ models in their predictive capabilities. We present the comparative analysis based on the RF models. It is important to note that the RF approach is an ensemble approach that is a predictive model, akin to using a weighted average of many nonlinear regressions (see, inter alia, [72]). However, any appropriate forecasting model (e.g., autoregressive integrated moving average (ARIMA($p, d, q$)), can also be used to compare the predictive performances of the covariates.

A RF model sorts the predictor space into a number of non-overlapping regions $R_1, R_2, \ldots, R_n$ and makes a top-down decision tree. A common dividing technique is recursive binary splitting process, where in each split it makes two regions $R_1 = \{X|X_i < k\}$ and $R_2 = \{X|X_i \geq k\}$ by considering all possible predictors $X_j$s and their corresponding cutpoint $k$ such that residual sum of squares (RSS) (Eq. (11)) becomes the lowest.

$$RSS = \sum_{x_i \in R_1(j,k)} (y_i - \hat{y}_{R_1})^2 + \sum_{x_i \in R_2(j,k)} (y_i - \hat{y}_{R_2})^2,$$

where $\hat{y}_{R_1}$ and $\hat{y}_{R_2}$ are the mean responses for the training observations in the region $R_1(j, k)$, and in $R_2(j, k)$, respectively. To improve the predictive accuracy, instead of fitting a single tree, the RF technique builds a number of decision trees and averages their individual predictions [73]. RF is a non-linear model (piece-wise linear). Therefore, if there is any nonlinear causality of $X$ to $AP$ and $V$, a RF model captures this causality.

We compare the predictive performance of a baseline model (Model $P_0$), which includes only the lagged values of the abnormal price, with other proposed models which additionally include a set of covariates. The covariates are selected based on their significant correlations and causalities. Table 2 represents a description of the five models we use in our analysis.

We consider the root mean squared error (RMSE) as measure of prediction error. The RMSE for abnormal price modeling can be defined as

$$RMSE = \sqrt{(1/n) \sum_{t=1}^{n} (y_t - \hat{y}_t)^2},$$

where $y_t$ is the test set of abnormal price (AP) and $\hat{y}_t$ is the corresponding predicted value. We calculate the percentage change in prediction error (RMSE) for a specific model in Table 2 with respect to model $P_0$ as

$$\Delta = (1 - \frac{\Psi(P_i)}{\Psi(P_0)}) \times 100%, \quad i = 1, \ldots, 4,$$

where $\Psi(P_i)$ and $\Psi(P_0)$ are the RMSE of model $P_0$ and model $P_i$, respectively. If $\Delta > 0$, the covariate ($X$) is said to improve prediction of $Y$. We compare the $\Delta$ for different models, calculated for varying prediction horizons.

3.3. Analysis of volatility

We evaluate the utility of COVID-19 cases and deaths, US county level spread of COVID-19, and Google searches in predicting stock market volatility. Let the conditional mean of log return of S&P 500 price ($r_t$) be given as

$$y_t = E(y_t|I_{t-1}) + \epsilon_t,$$

where $I_{t-1}$ is the information set at time $t - 1$, and $\epsilon_t$ is conditionally heteroskedastic error. We build two exponential GARCH (EGARCH ($p, q$)) models, Model 0 and Model X, where Model 0 is a standard EGARCH model with no explanatory variables, and Model X includes a set of explanatory variables:

$$\epsilon_t = \sigma_t \eta_t,$$

7
Fig. 4. Time plots of abnormal price (AP) and volatility (Vol) from January 13 2020 to May 29 2020.

Model 0: \[ \log_e (\sigma^2_t) = \omega_0 + \sum_{i=1}^{q} (\omega_i \eta_{t-j} + \gamma_j (|\eta_{t-j} - E| \eta_{t-j})) + \sum_{j=1}^{p} \tau_j \log_e (\sigma^2_{t-j}), \]

Model X: \[ \log_e (\sigma^2_t) = \omega_0 + \sum_{i=1}^{q} (\omega_i \eta_{t-j} + \gamma_j (|\eta_{t-j} - E| \eta_{t-j})) + \sum_{j=1}^{p} \tau_j \log_e (\sigma^2_{t-j}) + \Lambda X_t, \tag{14} \]

where \( \eta_t \sim \text{iid}(0,1), i = 1, 2, \ldots, q, j = 1, 2, \ldots, p \). [74–78].

We select a set of eight explanatory variables: \( X = [\text{US total deaths lag 1}, \text{US total deaths lag 2}, \# \text{Edges lag 1}, \# \text{Edges lag 2}, T_1 \text{ lag 1}, T_2 \text{ lag 1}, \text{“Covid 19” US lag 1}, \text{“Covid 19” US lag 2}] \) with \( \Lambda = [\lambda_1 \lambda_2 \ldots \lambda_8] \). All the explanatory variables are in the form of log returns. For simplicity we choose EGARCH (1,1) model. For EGARCH (1,1) with the assumption of \( \eta_t \sim \text{iid}(0,1) \) the two propose models (Eq. (14)) reduce to

Model 0: \[ \log_e (\sigma^2_t) = \omega_0 + \omega_i \eta_{t-j} + \gamma_j |\eta_{t-j} - E| \eta_{t-j}) + \tau_j \log_e (\sigma^2_{t-j}), \]

Model X: \[ \log_e (\sigma^2_t) = \omega_0 + \omega_i \eta_{t-j} + \gamma_j |\eta_{t-j} - E| \eta_{t-j}) + \tau_j \log_e (\sigma^2_{t-j}) + \sum_{l=1}^{8} \lambda_l x_{lt}. \tag{15} \]

The performances of the two models are compared based on their log likelihood, Akaike Information Criterion (AIC) and Bayesian information criterion (BIC).

4. Results

We investigate the effect of COVID-19 public health crisis on the stock market, in particular, on the S&P 500 index. We primarily focus on the impact of COVID-19 cases and deaths, local spread, and COVID-19 related Google searches on S&P 500. Fig. 4 shows the movements of abnormal S&P 500 price and volatility from January 13, 2020 to May 29, 2020. The top panel illustrates the precipitous drop of S&P 500 price (Eq. (1)), and historic high volatility (Eq. (2)) is depicted in the bottom panel.

We start our analysis with the Spearman’s rank correlation test. We calculate correlations between the daily abnormal S&P 500 closing price, AP and the daily COVID-19 cases and deaths, and daily occurrences of higher order structures in the spread network at different time lags. For example, at lag 1 we compute correlation of AP at day \( t \) with COVID-19 cases and deaths, and higher order network structures, all at day \( t - 1 \). Fig. 5(a) shows the box plots which combine correlations between each COVID-19 cases and deaths variable and AP at different lag. Here we build two box plots at each lag: one for COVID-19 cases and deaths in the US (four variables), and another for COVID-19 cases and deaths in the World (four variables). Similarly, Fig. 5(b) represents the box plots that combined correlations between each eleven local spread variable and AP at different lag.

We find that there exists significant (negative) correlation between COVID-19 cases and deaths in US and abnormal S&P 500 in all six lags, \( lag = 1, 2, \ldots, 6 \). However, there is no significant correlation between COVID-19 cases and deaths in the entire world and abnormal S&P 500 (\( p \)-value > 0.05) in any lag (see Table 7 in Appendix). We also find that all the local spread variables are significantly (negative) correlated (\( p \)-value < 0.05) with abnormal S&P 500 in every
Fig. 5. (Spearman) Correlations between Covid-19 and abnormal S&P 500. Correlations of eight COVID-19 variables in each lags are summarized in a box plot.

Table 3
Summary of G-causality analysis of COVID-19 cases and deaths on abnormal S&P 500 ($y$) on different lag effects (day). $P$ and $Vol$ denote significance in price and volatility, respectively. Blank space implies no significance. Confidence level is 90%.

| Causality               | Lag |
|------------------------|-----|
|                        | 1   | 2   | 3   | 4   | 5   | 6   | 7   |
| US total cases → $y$   | –   | –   | –   | $Vol$ | –   | $Vol$ | –   |
| US total deaths → $y$  | P/$Vol$ | $Vol$ | P/$Vol$ | $Vol$ | P/$Vol$ | $Vol$ | $Vol$ |
| US new cases → $y$     | P/$Vol$ | –   | –   | $Vol$ | $P$  | P/$Vol$ | $P$  |
| US new deaths → $y$    | –   | –   | –   | –   | –   | –   | –   |
| World total cases → $y$| –   | –   | –   | –   | –   | –   | –   |
| World total deaths → $y$| –   | –   | –   | –   | –   | –   | –   |
| World new cases → $y$  | –   | –   | –   | –   | –   | –   | –   |
| World new deaths → $y$ | –   | $P$ | $P$ | $P$ | $Vol$ | $Vol$ | $Vol$ |

$lag = 1, 2, \ldots, 6$. That is, US county level spread of COVID-19 adversely affects the price of S&P 500. However, it is anticipated that the strength of correlations of local spread variables will gradually decrease in higher lags, which is also reflected in Fig. 5(b). Some of the COVID-19 related Google searches, e.g., “Covid-19” in US and “Corona” in world are also significantly correlated ($p$-value > 0.1) with abnormal S&P 500 in different lags (Table 12 in Appendix).

The link with the domestic variables is similar to the channel proposed by Cao et al. [20], but in contrast with their results and with the results presented in Onali [17], we find that international variables (in our case deaths in the whole world) do not play a significant role. However, it is important to note that our global variables are measured differently. Even though we use very different methodological approach from the approach used by Rubban et al. [23], the fact that local spreads have strong predictive power for returns parallels their message that deficiencies in the health sector can directly cause negative effects on the stock market.

We also investigate the potential impact of COVID-19 cases and deaths, its local spread, and related Google searches on S&P 500 price formation and risk, i.e., volatility. Tables 3 and 4 present summaries of the Granger causality tests for predictive utility of COVID-19 cases and deaths, and county level local spread, respectively. Here the direction of causality is denoted by $\rightarrow$.

We find that US total new cases and US total death have significant predictive impacts on price and volatility. US total number of cases have predictive relationship only with volatility in few lags. Among world COVID-19 cases and deaths only total new deaths have causality on price and volatility. Almost all the local spread variables have predictive impact on price, but none of them except # Edges at lag 1 have causality on volatility. That is, county level spread of COVID-19 significantly influence abnormal price formation, but, surprisingly, they do not have causal linkage with the volatility. Table 10 in Appendix shows that a number of Google search variables have causality effects on abnormal price. However, only “Coronavirus” US and “Covid 19” US have predictive impacts on volatility at very few lags. The link between volatility and national factors is similar to the findings of [17]. One of the key novel findings of our study is that while volatility is mostly affected by national factors, the price level responds to local spread and to sentiment, where “local” is defined on a much more granular level than in previous studies.
Table 4
Summary of G-causality analysis of COVID-19 spread on abnormal S&P 500 (y) on different lag effects (day). P and Vol denote significance in price and volatility, respectively. Blank space implies no significance. Confidence level is 90%.

| Causality          | Lag | 1     | 2     | 3     | 4     | 5     | 6     | 7     |
|--------------------|-----|-------|-------|-------|-------|-------|-------|-------|
| # Edges → y       |     | Vol   | P     | P     | P     | P     | P     | P     |
| GC → y            |     | P     | P     | P     | P     | P     | P     | P     |
| T1 → y            |     | P     | P     | –     | –     | –     | –     | –     |
| T2 → y            |     | P     | P     | P     | P     | P     | P     | P     |
| V1 → y            |     | –     | –     | P     | –     | –     | –     | –     |
| V2 → y            |     | P     | P     | –     | P     | –     | –     | P     |
| V3 → y            |     | –     | –     | –     | –     | –     | –     | –     |
| V4 → y            |     | –     | P     | P     | P     | P     | P     | P     |
| V5 → y            |     | –     | P     | P     | P     | P     | P     | P     |
| V6 → y            |     | –     | –     | –     | –     | –     | –     | –     |
| Total # V → y     |     | –     | –     | P     | –     | –     | –     | –     |

Table 5
Predictive utilities (Δ) of models in Table 2 over the baseline model (Model P0) for different prediction horizons.

| h     | Model P1 | Model P2 | Model P3 | Model P4 |
|-------|----------|----------|----------|----------|
| 1     | −0.411   | −5.305   | 7.219    | 2.086    |
| 2     | 0.171    | −1.257   | 2.279    | 0.042    |
| 3     | 1.549    | 3.242    | 3.477    | 2.397    |
| 4     | 1.463    | 3.579    | 3.368    | 2.672    |
| 5     | 1.410    | 3.843    | 2.718    | 2.922    |
| 6     | 0.962    | 3.898    | 2.718    | 3.107    |

Fig. 6. Abnormal price prediction for March 2020 to May 2020 with 1, and 2 day horizons.

Now we turn our analysis to compare the predictive performance of models described in Table 2. Table 5 presents prediction errors (based on Eq. (12)) calculated for varying prediction horizons h = 1, 2, . . . , 6. We find that, for short term forecasting horizons (h = 1, 2, and 3) model P3, which is based on Google search variables yields more accurate performance. For longer term forecasting horizons (h = 4, 5, and 6), model P2 containing information from local spread delivers the most competitive results, followed by model P4, which contains information from COVID-19 deaths, local spread, and Google searches.

Fig. 6 represents a comparison of the observed data with fitted values from baseline model (model P0) and four other models, i.e., model P1, P2, P3, and P4. For 1 day horizon model P1 yield a noticeably higher predictive accuracy followed by model P4. For 2 day horizon, although it is expected that the prediction performances of all models deteriorates compare to their performances for 1 day horizon, model P3 again delivers the best prediction accuracy.

We now evaluate the influence of COVID-19 cases and deaths, US county level spread of COVID-19, and Google searches on S&P 500 volatility. A comparison of the two EGARCH models, Model 0 and Model X (Eq. (15)), including the estimated parameters of the explanatory variables for Model X are presented in Table 6. All EGARCH coefficients except the constant term (ω0) are statistically significant in both models. However, the coefficients estimates of all the covariates in Model X are not statistically significant.

We also examine the goodness of fit of the two models by comparing their log likelihood, Akaike Information Criterion (AIC) and Bayesian information criterion (BIC). We find that Model 0 tends to describe the S&P 500 volatility more
Table 6
Estimates of EGARCH models for S&P 500 price volatility. **p < 0.01, *p < 0.05, *p < 0.1.

| Parameter | Model X Coef. | t value | Coef. | t value |
|-----------|---------------|---------|-------|---------|
| $\omega_0$ | -0.611 | -1.287 | -0.392 | -1.394 |
| $\omega$ | -0.517 | -3.249*** | -0.484 | -3.443*** |
| $\gamma$ | 0.514 | 2.387*** | 0.513 | 2.871*** |
| $\tau$ | 0.937 | 16.192*** | 0.955 | 2.871*** |
| US total deaths lag 1 ($\lambda_1$) | -0.867 | -0.864 |
| US total deaths lag 2 ($\lambda_2$) | -0.558 | -0.609 |
| # Edges lag 1 ($\lambda_3$) | -0.260 | -1.235 |
| # Edges lag 2 ($\lambda_4$) | 0.050 | 0.221 |
| $T_2$ lag 1 ($\lambda_5$) | -0.514 | -0.939 |
| $T_2$ lag 2 ($\lambda_6$) | -0.067 | -0.120 |
| “Covid 19” US lag 1 ($\lambda_7$) | -0.132 | -0.167 |
| “Covid 19” US lag 2 ($\lambda_8$) | -0.476 | -0.621 |
| Log-likelihood | 214.049 | 218.258 |
| AIC | -4.540 | -4.709 |
| BIC | -4.205 | -4.599 |

Fig. 7. Time plots of $AP$ and $Vol$ from January 13 2020 to May 29 2020.

accurately than the volatility model with covariates, Model X. That is, COVID-19 cases and deaths, its local spread and Google searches do not significantly influence the S&P 500 volatility. Fig. 7 also suggests that Model 0 captures the spikes of the price returns more accurately than Model X.

COVID-19 related factors do not seem to play a role when it comes to explaining the dynamics of the volatility. One potential explanation for this is that the movements in the volatility are driven by national-level sentiment about policy or about policy uncertainty. As a robustness check, we perform an experiment where we explore the correlations and Granger causality between Google trend searches for macroeconomic policy variables. Because the data for the COVID-19 spread is available at daily frequency, we use daily data in all of our specifications. Therefore, we proxy for investor sentiment about policy by using Google Trend searches and the policy uncertainty index from Baker et al. [63] rather than higher frequency 30 min windows around policy announcements as in Cox et al. [24]. The results are reported in Appendix. Our results are very similar to the results for the benchmark specifications with the volatility being affected by national factors. While sentiment about policy is correlated with abnormal prices, we only find Granger causality in several cases related to fiscal policy searches, and only at higher lags. On the other hand, sentiment about policy Granger-causes volatility at all lags.

5. Conclusion and discussion

The aim of this paper is to evaluate whether COVID-19 cases and deaths, local spread of COVID-19, and Google search activity explain and predict US stock market plunge in the spring of 2020. We develop a modeling framework that systematically evaluates the correlation – causality – predictive utility of each of the COVID-19 related features
on stock decline and stock volatility. In order to quantify local spread of COVID-19 we construct a temporal spread network and study the dynamics of higher order network structures as a measure of local spread. We find that COVID-19 cases and deaths, its local spread, and Google search activities related to COVID-19 have contemporary relationships and predictive abilities on abnormal stock prices. Our results indicate that COVID-19 cases and deaths, and its local spread not only unprecedentedly disrupt economic activity and cause a collapse in demand for different goods but also they make investors panic and increase their anxiety. The anxiety is also reflected in Google search intensity for COVID-19. These shocks affect investment decisions and the subsequent stock price dynamics. On the other hand, very few COVID-19 variables have causal relationship on volatility. However, standard EGARCH models for the volatility show that COVID-19 cases and deaths, its local spread, and Google search volumes do not have impact on volatility. Different forms for the volatility measure \cite{37,38,79} lead to the same conclusions.

Overall, the volatility is mostly affected by national factors and incorporating higher-order information about local spread does not significantly improve the forecasting performance of the models. However, the local spread are significantly linked to abnormal price returns. Furthermore, incorporating information about the local spread significantly improves the predictive performance of the models for the abnormal price level.

Even though our contribution is methodological, the implications for investors are quite direct: the local spreads provided strong predictive improvements for stock prices and they should be included in predictive models. More broadly, along the same line as Hernández-Murillo and Owyang \cite{29}, we find that incorporating information about local variables improves the predictive power for models for national variables. One limitation of our study is that we focus only on the effects of the early stages of the pandemic on the US stock market. In future extensions of this work we would like to extend our model to consider the effects on effect across US equity sectors and across local US labor markets over a longer time horizon, which would allow us to disentangle the sentiment channel. Furthermore, taking into account the fact that the COVID-19 pandemic started less than 2 years, our study cannot provide long-run guidance for investors at this point. However, despite this limitation, our findings are quite conclusive about the predictive importance of local spreads.

Our results also have implications for policymakers when it comes to public health, monetary, and fiscal policy measures. Given the direct link between local spreads and prices, our main suggestion parallels the findings from Shehzad et al. \cite{22} where deficiency in the health sector can damage the stock market more than (and in addition) to the economic crisis. Bolstering the public health sector and controlling local spread may have positive effects on the market. When it comes to monetary policy, public health policies and the Federal reserve's commitment to transparency, more relaxed inflation goals, and commitment to financial stability can work in conjunction to further shore up financial markets. Furthermore, as pointed by Bouri et al. \cite{25}, financial markets were more interconnected during the COVID-19 crisis. Both public health policies that stabilize spread and timely monetary policy actions that are aimed at directly stabilizing one market and that take into account potential spillovers may have further indirect benefits by preventing contagion. While local spreads did not have strong predictive power for the overall market volatility, national COVID-19 variables did, indicating that public health, monetary, and fiscal policy can work in conjunction to stabilize the economy. Fiscal stimulus can have large positive effects on the real sector in times of elevated uncertainty, in times of elevated economic slack, and when the economy is close to the zero lower bound \cite{80–82}. If stimulus measures are combined with focused support to contain local spreads, this may have positive and stabilizing effects both on the real and on the financial sector.

Our main contributions are twofold. First, we leverage mathematical tools to construct a novel measure of local spread that goes beyond simply using local prevalence. This tool can be applied in other empirical applications for modeling other diseases or financial market contagion. Much more importantly, we show that including the local spread variables significantly improves the performance of the predictive models for abnormal prices.

CRediT authorship contribution statement

Asim K. Dey: Conceptualization, Methodology, Investigation, Writing – original draft. G.M. Toufigul Hoque: Data curation, Software. Kumer P. Das: Methodology, Supervision, Writing – review & editing. Irina Panovska: Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Dey has been partially supported by NSF DMS, United States 2027793. We would like to thank four reviewers for their constructive comments and suggestions.

Appendix

(see Tables 7–12).
**Table 7**
Spearman correlations between COVID-19 cases and abnormal S&P 500. Blue color indicates significant correlation (p-values < 0.05), while black color represents non-significant correlation (p-values > 0.05).

| Lag | 0    | 1    | 2    | 3    | 4    | 5    | 6    |
|-----|------|------|------|------|------|------|------|
| US total cases | -0.45 | -0.48 | -0.53 | -0.59 | -0.60 | -0.59 | -0.55 |
| US total deaths | -0.75 | -0.76 | -0.78 | -0.82 | -0.83 | -0.82 | -0.78 |
| US new cases | -0.37 | -0.41 | -0.42 | -0.45 | -0.48 | -0.48 | -0.40 |
| US new deaths | -0.37 | -0.34 | -0.36 | -0.40 | -0.45 | -0.45 | -0.39 |
| World total cases | -0.04 | -0.05 | -0.01 | 0.00  | -0.03 | -0.06 | -0.07 |
| World total deaths | -0.08 | -0.11 | -0.11 | -0.18 | -0.20 | -0.19 | -0.17 |
| World new cases | -0.14 | -0.18 | -0.13 | -0.15 | -0.17 | -0.18 | -0.18 |
| World new deaths | -0.11 | -0.09 | -0.10 | -0.17 | -0.21 | -0.17 | -0.14 |

**Table 8**
Spearman correlations between Local spread variables and abnormal S&P 500. Blue color indicates significant correlation (p-values < 0.05). A non-significant correlation (p-values > 0.05) is presented by black color.

| Lag | 0    | 1    | 2    | 3    | 4    | 5    | 6    |
|-----|------|------|------|------|------|------|------|
| Edge | 0.60 | -0.58 | -0.63 | -0.64 | -0.61 | -0.60 | -0.57 |
| GC  | 0.61 | -0.43 | -0.50 | -0.50 | -0.50 | -0.51 | -0.49 |
| T1  | 0.95 | -0.35 | -0.35 | -0.34 | -0.31 | -0.30 | -0.30 |
| T2  | 0.90 | -0.35 | -0.36 | -0.35 | -0.34 | -0.32 | -0.32 |
| M1  | 0.82 | -0.22 | -0.23 | -0.22 | -0.20 | -0.19 | -0.20 |
| M2  | 0.93 | -0.31 | -0.30 | -0.30 | -0.28 | -0.27 | -0.27 |
| M3  | 0.98 | -0.35 | -0.35 | -0.35 | -0.32 | -0.31 | -0.30 |
| M4  | 0.84 | -0.27 | -0.26 | -0.25 | -0.25 | -0.22 | -0.22 |
| M5  | 0.84 | -0.36 | -0.35 | -0.35 | -0.33 | -0.32 | -0.31 |
| M6  | 0.91 | -0.40 | -0.40 | -0.38 | -0.38 | -0.36 | -0.35 |
| Total | 0.93 | -0.39 | -0.39 | -0.38 | -0.37 | -0.35 | -0.33 |

**Table 9**
Spearman correlations between google trend and abnormal S&P 500. A significant correlation (p-values < 0.05) is represented by blue color, while black color indicates a non-significant correlation (p-values > 0.05).

| Lag | 0    | 1    | 2    | 3    | 4    | 5    | 6    |
|-----|------|------|------|------|------|------|------|
| “Coronavirus” US | 0.02 | -0.03 | -0.10 | -0.11 | -0.08 | -0.07 | -0.10 |
| “Corona” US | 0.25 | 0.18  | 0.12  | 0.11  | 0.09  | 0.05  | 0.02  |
| “Covid-19” US | -0.26 | -0.31 | -0.28 | -0.28 | -0.29 | -0.29 | -0.32 |
| “Covid 19” US | -0.09 | -0.15 | -0.17 | -0.21 | -0.25 | -0.29 | -0.30 |
| “Coronavirus” World | 0.10 | 0.08  | 0.1   | -0.03 | -0.05 | -0.08 | -0.06 |
| “Corona” World | 0.26 | 0.22  | 0.16  | 0.11  | 0.06  | 0.02  | 0.00  |
| “Covid-19” World | -0.04 | -0.07 | -0.08 | -0.10 | -0.11 | -0.12 | -0.14 |
| “Covid 19” World | -0.11 | -0.11 | -0.12 | -0.15 | -0.15 | -0.21 | -0.25 |

**Table 10**
G-causality analysis of Google searches on abnormal S&P (y) on different lag effects (day). P and Vol denote significance in price and volatility, respectively. Blank space implies no significance. Confidence level is 90%.

| Causality | Lag | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|-----|---|---|---|---|---|---|---|
| “Coronavirus” US --> y | - | - | - | - | - | - | - |
| “Covid-19” US --> y | P | P | P | P | P | P | P |
| “Covid 19” US --> y | Vol | P | P | P | P | - | - |
| “Covid - 19” US --> y | - | P | - | - | - | - | - |
| “Coronavirus” World --> y | - | - | P | - | - | - | - |
| “Covid-19” World --> y | P | P | P | P | P | P | P |
| “Covid 19” World --> y | - | - | - | - | - | - | - |
| “Covid - 19” World --> y | P | P | P | P | - | - | - |
Table 11
Spearman correlations between Economic Policy Uncertainty (EPU) Index and google trend in US related to economic policy, and abnormal S&P 500. A significant correlation (p-values < 0.05) is represented by blue color, while black color indicates a non-significant correlation (p-values > 0.05).

| Lag | EPU | “Unemployment benefit” | “Stimulus package” | “Coronavirus stimulus” | “Stimulus” | “Stimulus check” | “irs stimulus” |
|-----|-----|------------------------|--------------------|------------------------|------------|-----------------|---------------|
| 0   | -0.276 | -0.382 | -0.177 | -0.192 | -0.159 | -0.058 | -0.289 |
| 1   | -0.277 | -0.388 | -0.150 | -0.194 | -0.155 | -0.0374 | -0.282 |
| 2   | -0.300 | -0.357 | -0.137 | -0.189 | -0.171 | -0.073 | -0.282 |
| 3   | -0.280 | -0.340 | -0.151 | -0.244 | -0.162 | -0.058 | -0.323 |
| 4   | -0.268 | -0.285 | -0.164 | -0.249 | -0.171 | -0.071 | -0.314 |
| 5   | -0.308 | -0.242 | -0.148 | -0.223 | -0.162 | -0.057 | -0.305 |

Table 12
G-causality analysis of Economic Policy Uncertainty (EPU) Index and google trend in US related to economic policy on abnormal S&P (y) on different lag effects (day). P and Vol denote significance in price and volatility, respectively. Blank space implies no significance. Confidence level is 90%.

| Lag | EPU | “Unemployment benefit” → y | “Stimulus package” → y | “Coronavirus stimulus” → y | “Stimulus” → y | “Stimulus check” → y | “irs stimulus” → y |
|-----|-----|-----------------------------|------------------------|---------------------------|----------------|---------------------|------------------|
| 1   | -    | -                           | -                      | -                         | -              | -                   | -                |
| 2   | -    | -                           | -                      | -                         | -              | -                   | -                |
| 3   | -    | -                           | -                      | -                         | -              | -                   | -                |
| 4   | -    | -                           | -                      | -                         | -              | -                   | -                |
| 5   | -    | -                           | -                      | -                         | -              | -                   | -                |
| 6   | -    | -                           | -                      | -                         | -              | -                   | -                |
| 7   | -    | -                           | -                      | -                         | -              | -                   | -                |

References

[1] J.A. Cagle, Natural disasters, insurer stock prices, and market discrimination: The case of hurricane hugo, J. Insur. Issues 19 (1) (1996) 53–68.
[2] E. Cavallo, I. Noy, The Economics of Natural Disasters - A Survey, Working Papers 200919, University of Hawaii at Manoa, Department of Economics, 2009.
[3] A. Worthington, A. Valadkhani, Measuring the impact of natural disasters on capital markets: an empirical application using intervention analysis, Appl. Econ. 36 (19) (2004) 2177–2186, http://dx.doi.org/10.1080/0003684042000282489.
[4] A.C. Worthington, The impact of natural events and disasters on the Australian stock market: a GARCH-m analysis of storms, floods, cyclones, earthquakes and bushfires, Glob. Bus. Econ. Rev. 10 (2008) 1–10.
[5] L. Shan, S.X. Gong, Investor sentiment and stock returns: Wenchuan Earthquake, Finance Res. Lett. 9 (1) (2012) 36–47.
[6] R. Hudson, A. Urquhart, War and stock markets: The effect of World War Two on the British stock market, Int. Rev. Financ. Anal. 40 (2015) 166–177.
[7] G. Schneider, V.E. Troeger, War and the world economy: Stock market reactions to international conflicts, J. Conflict Resolution 50 (5) (2006) 623–645.
[8] F. Chau, R. Deemosak, J. Wang, Political uncertainty and stock market volatility in the Middle East and North African (MENA) countries, J. Int. Financ. Mark. Inst. Money 28 (2014) 1–19.
[9] M.C. Beaulieu, J.C. Cosset, N. Essaddam, Political uncertainty and stock market returns: evidence from the 1995 Quebec referendum, Can. J. Econ./Rev. Can. écon. 39 (2) (2006) 621–642.
[10] T. Huynh, T. Burggraf, If worst comes to worst: Co-movement of global stock markets in the US-China trade war, SSRN Electron. J. (2019).
[11] S. Nippani, K.M. Washer, SARS: A non-event for affected countries’ stock markets? Appl. Financial Econ. 14 (15) (2004) 1105–1110.
[12] A. Siu, Y.C.R. Wong, Economic impact of SARS: The case of Hong Kong, Asian Econ. Pap. 3 (1) (2004) 62–83.
[13] J.W. Lee, W.J. McBibin, Globalization and disease: The case of SARS, Asian Econ. Pap. 3 (1) (2004) 113–131.
[14] R. Ichev, M. Mariné, Stock prices and geographic proximity of information: Evidence from the Ebola outbreak, Int. Rev. Financ. Anal. 56 (2018) 153–166.
[15] M. Nicola, Z. Alsafi, C. Soroabi, A. Kerwan, A. Al-Jabir, C. Isfotidis, M. Agha, R. Agha, The socio-economic implications of the coronavirus and COVID-19 pandemic: A review, Int. J. Surgery 78 (2020).
[16] S.R. Baker, N. Bloom, S.J. Davis, K.J. Kost, M.C. Sammon, T. Viratyosin, The unprecedented stock market impact of COVID-19, Working Paper Series 26945, National Bureau of Economic Research, 2020.
[17] E. Onali, Covid-19 and stock market volatility, 2020, http://dx.doi.org/10.2139/ssrn.3571453.
[18] A. Zaremba, R. Kizys, D.Y. Aharon, E. Demir, Infected markets: Novel coronavirus, government interventions, and stock return volatility around the globe, Finance Res. Lett. 35 (2020) 101597.
[19] K. Arias-Calhuari, F. Alonso-Marroquin, M.N. Najaﬁ, M. Harré, Methods for forecasting the effect of exogenous risks on stock markets, Physica A 568 (2021) 125587.
[20] K.H. Cao, Q. Li, Y. Liu, C.K. Woo, Covid-19's adverse effects on a stock market index, Appl. Econ. Lett. (2020) 1–5.
[21] A. Wagner, What the stock market tells us about the post-COVID-19 world, Nat. Hum. Behav. 4 (2020) http://dx.doi.org/10.1038/s41562-020-0869-y.
[22] K. Shehzad, L. Xiaoqing, M. Arif, K.U. Rehman, M. Ilyas, Investigating the psychology of financial markets during COVID-19 era: A case study of the US and European markets, Front. Psychol. 11 (2020) 1924.
[23] G. Rubbanby, A.A. Khalid, M. Umar, M. Nawazish, European stock markets’ response to Covid-19, lockdowns, government stringency and central bank interventions, Working Paper Series 3758227, SSRN, 2020.
[24] J. Cox, D. Greenwald, S.C. Ludvigson, What Explains the COVID-19 Stock Market, Working Paper Series 27784, National Bureau of Economic Research, 2020.
[74] D.B. Nelson, Conditional heteroskedasticity in asset returns: A new approach, Econometrica 59 (2) (1991) 347–370.
[75] M. McAleer, C.M. Hafner, A one line derivation of EGARCH, Econometrics 2 (2) (2014) 92–97.
[76] C.L. Chang, M. McAleer, The correct regularity condition and interpretation of asymmetry in EGARCH, Econom. Lett. 161 (2017) 52–55.
[77] G.G. Martinet, M. McAleer, On the invertibility of EGARCH(p, q), Econometric Rev. 37 (8) (2018) 824–849.
[78] T. Bollerslev, A.J. Patton, R. Quaedvlieg, Multivariate leverage effects and realized semicovariance GARCH models, J. Econometrics 217 (2) (2020) 411–430, Nonlinear Financial Econometrics.
[79] P. Molnár, Properties of range-based volatility estimators, Int. Rev. Financ. Anal. 23 (2012) 20–29.
[80] S.M. Fazzari, J. Morley, I. Panovska, State-dependent effects of fiscal policy, Stud. Nonlinear Dyn. Econom. 19 (3) (2015) 285–315.
[81] G. Caggiano, E. Castelnovo, V. Colombo, G. Nodari, Estimating fiscal multipliers: News from a non-linear world, Econ. J. 125 (2015) 746–776.
[82] S.M. Fazzari, J. Morley, I. Panovska, When is discretionary fiscal policy effective? Stud. Nonlinear Dyn. Econom. (2020).