Analyzing airlines stock price volatility during COVID-19 pandemic through internet search data

Soudeep Deb

Indian Institute of Management, Bangalore
Bannerghatta Main Rd, Bilekahalli
Bengaluru, Karnataka 560076, India.
Email: soudeep@iimb.ac.in

Abstract

Recent Coronavirus pandemic has prompted many regulations which are affecting the stock market. Especially because of lockdown policies across the world, the airlines industry is suffering. We analyze the stock price movements of three major airlines companies using a new approach which leverages a measure of internet concern on different topics. In this approach, Twitter data and Google Trends are used to create a set of predictors which then leads to an appropriately modified GARCH model. In the analysis, first we show that the ongoing pandemic has an unprecedented severe effect. Then, the proposed model is used to analyze and forecast stock price volatility of the airlines companies. The findings establish that our approach can successfully use the effects of internet concern for different topics on the movement of stock price index and provide good forecasting accuracy. Model confidence set procedure further shows that the short-term volatility forecasts are more accurate for this method than other candidate models. Thus, it can be used to understand the stock market during a pandemic in a better way. Further, the proposed approach is attractive and flexible, and can be extended to other related problems as well.

Keywords: American Airlines, COVID-19, Delta Airlines, Google Trends, Twitter data, United Airlines

1   Introduction

The year 2020 has been hit severely by a pandemic of a new kind, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-COV 2). Presumably, it started in December of 2019, when a cluster of cases resembling viral pneumonia were associated with a point source of spread from a fish market in Wuhan, Hubei, China (Shereen et al. (2020)). It has spread to more than 180 countries ever since and as of 1st May, 2020, it has already affected more than 3.4 million people, while the number of death tolls stands at more than quarter of a million worldwide. The outbreak was declared a Public Health Emergency of International Concern by the World Health Organization (WHO) on 30th January 2020. WHO named the disease as COVID-19 on 11th February, 2020. The reader is referred to Sohrabi et al. (2020) for a detailed review of the disease while Deb and Majumdar (2020), Roy and Karmakar (2020), Wu et al. (2020) and the references therein are some studies on understanding and predicting the number of cases affected by the virus.

Given that the disease spreads through cough droplets and is very contagious, most countries adopted complete or partial lockdown approaches to mitigate the spread of the disease. Many other policies, for example holiday extension, school and university closure, work from home system, hospitalisation and home quarantine, have also been adopted by different countries. Lin et al. (2020) researched on the effect of such individual reaction and governmental actions on the outbreak. Chinazzi et al. (2020), meanwhile, studied the effect of travel restrictions on the same. These papers confirm that such measures have helped contain the pandemic to some extent. However, an immediate consequence was the effect on economy. As Enserink and Kupferschmidt (2020) pointed out, long lockdowns to slow an epidemic can have catastrophic economic impacts and may thereby devastate public health as well. To that end, it is of utmost importance to understand the immediate and long term effects of COVID-19 on different industries. Guerrieri et al. (2020) and McKibbin and Fernando (2020) worth reading in this regard.

In this paper, our focus is on airline industry which is expected to suffer heavily due to the pandemic. Historically, it has been seen that the whole transportation industry is highly subject to the impacts of system risks, which can be triggered by a large number of external factors beyond control and the corresponding damages are inevitable. These factors include financial events (recession, fuel price change etc), natural calamities (hurricane, tsunami, polar vortex etc) or man-made disasters (war, terrorist attack etc). A pandemic is clearly one such factor which can significantly affect the airlines industry. It is no secret that globalization has helped the airlines industry grow considerably over the last couple of decades. But, with COVID-19 gripping over the whole world, tourism has taken a halt almost everywhere. Fernandes (2020) discussed how various countries have been suffering from lack of tourism during the last few months. In another recent study, Iacus et al. (2020) worked with the historical air traffic data and projected the volume of the same under different assumptions on the ongoing pandemic. It is evident that all airlines companies across the world are incurring huge losses as a consequence. An immediate result was visible on the stock price index. Starting from the end of January 2020, the stock price for most airlines dropped and a higher than usual level of volatility was also observed. This erratic behavior of the stock market was unprecedented and was not specific to only the airlines industry, cf. Baker et al. (2020). In order to to understand how severe the effect is at the moment and to what extent even the big airlines companies can suffer, in this paper we develop a method which uses the internet concern of certain topics to better capture and forecast the effect of a pandemic on the stock price volatility. Throughout the paper, we work with the data for three major airlines of the United States of America (USA), namely American Airlines, Delta Airlines, and United Airlines. We however emphasize that the proposed method can easily be extended to analyze other airlines of interest to the reader. Hereafter, the above three airlines will be abbreviated as AA, DA and UA, respectively.

The main takeaways from the paper lie both in the modeling approach and the results. Generalized autoregressive conditional heteroskedastic models (see Section 3.2 for more discussions) and its variants are most popular in forecasting stock price volatility. However, the standard approaches can analyze the effect of a pandemic only by introducing related regressors (for example, number of COVID-19 cases,

lockdown policies, travel restrictions etc). In comparison, we provide a different and more interesting way to understand the effect of a pandemic on the volatility of airlines stock price index. Here, we focus on public reaction to different news and announcements. We combine Twitter data and Google search volumes to quantify the public interest, and show how that can be used in modeling airlines stock price returns. While somewhat similar approaches were taken to forecast crude oil prices (Yu et al. (2019)), to the best of our knowledge, such method has never been developed in the context of aviation industry. Further, we find that the model works better for forecasting short-term volatility for the stock price returns. Naturally, it has the potential to be adapted and used in other applications as well.

The rest of the paper is organized as follows. Descriptions of the data and some exploratory analysis are provided in Section 2. The methods are outlined in Section 3 while in Section 4, we discuss the main results of the study. Finally, concluding remarks, limitations of the study and future scopes are listed in Section 5. All of our calculations are carried out in RStudio version 1.2.5033, coupled with R version 3.6.2.

2 Preliminary analysis

2.1 Data

As mentioned earlier, throughout this paper, stock price indexes of AA, DA and UA are considered as the primary variables. We take the closing price for each available day, starting from 1st January, 2015 to 30th April, 2020. This data is obtained from Yahoo Finance. For price $P_{s,t}$ of stock s at time t, we compute the log-return as

$$R_{s,t} = 100 \left(\log P_{s,t} - \log P_{s,t-1} \right). \tag{2.1}$$

In Figure 1, the daily series of the returns for the three airlines stocks are presented. Dates for the two epidemic periods are also marked in the same graph. It is evident that the return values fluctuated a lot following the pandemic announcement about COVID-19, contrary to what was observed in the earlier pandemic. This is more clearly visible on the right panel of the figure, where we see the return series only in 2020. In the following discussion, we will quantify the daily volatility by taking absolute returns. Also, since we are primarily interested in short term volatility forecasting, throughout this paper, we keep last three weeks' data (after 12th April, 2020) as the test set and use everything else as the training set. That results in 1327 days worth of data in the training set.

Next, we take a look at the numbers for the COVID-19 outbreak in the USA. In addition, just for comparison, we include the 2017-18 flu season which was the last time USA suffered extensively due to an epidemic. It was widespread and as many as 32 states had high activity. Almost a million people were affected while in excess of sixty thousand people passed away. In comparison, COVID-19 has so far affected more than one million people and has claimed almost sixty three thousand lives in the USA. An overview of these two epidemics is presented in Table 1. The COVID-19 data are obtained from JHU-CSSE (2020) and the time range is from 22nd January, 2020 to 30th April, 2020.

Table 1: Comparison of 2017-18 flu season to the COVID-19 pandemic in the USA. All values are as of 1st May, 2020.

Disease	2017-18 Flu	COVID-19
Total cases	959,000	1,062,605
Total deaths	61,099	62,380
Fatality rate	6%	6%
Start date of outbreak	15th November, 2017	11th February, 2020

In the above table, 15th November, 2017 is taken as the start of the outbreak for the 2017-18 flu season as the influenza-like-illness activity began to increase around that time. The outbreak reached an
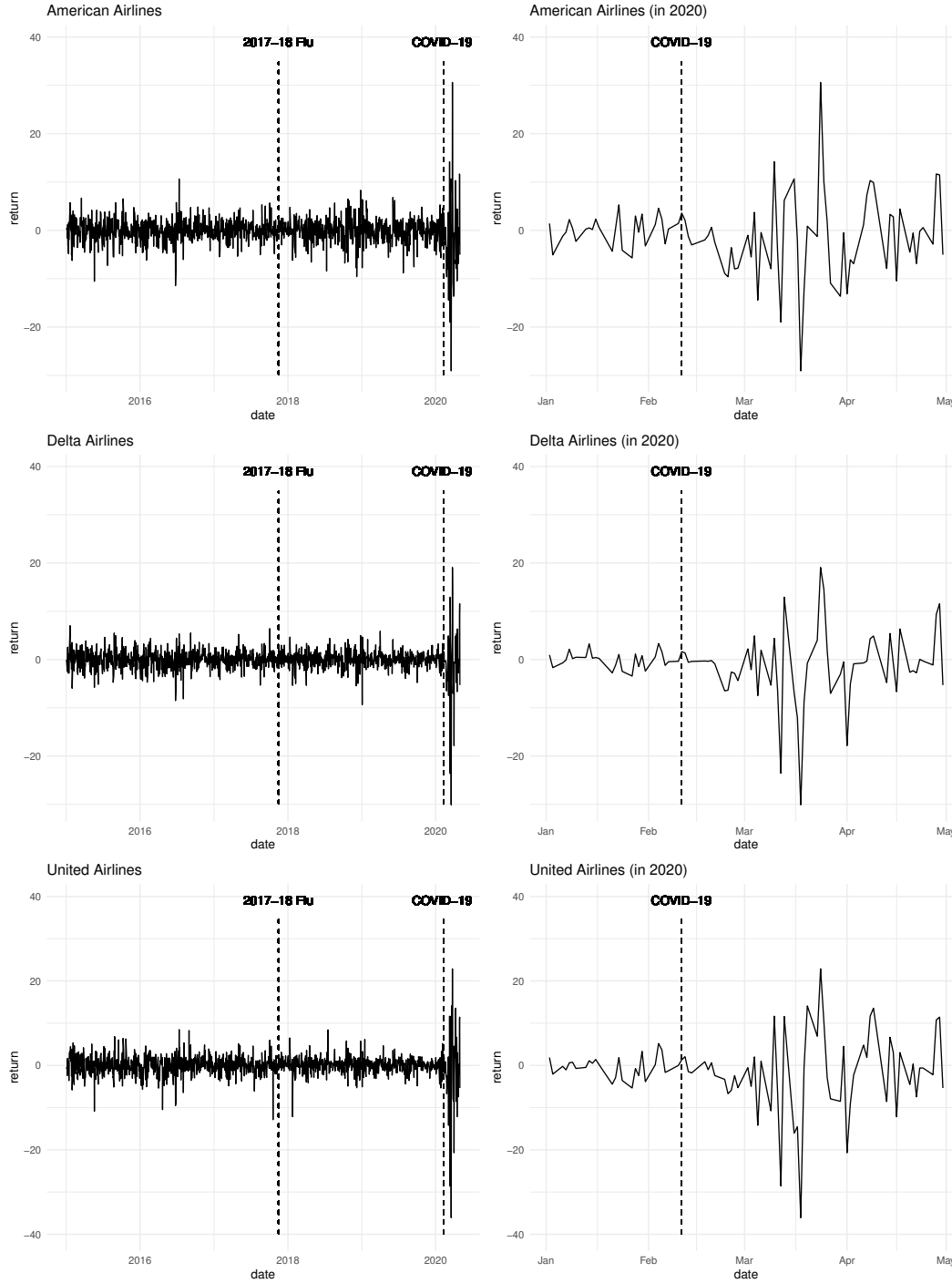
Figure 1: Time series of daily returns for the stock price index of three airlines. Left panel shows the plots since 2014, and the right panel shows the plots for the data from only 2020.

extended period of high activity during January and February across the whole country, and remained at that high level through the end of March. On the other hand, WHO made a declaration about COVID-19 on 11th February, 2020. The number of cases in the USA started to increase rapidly from around that time and it has not yet showed any sign of decreasing. Figure 2 shows the daily series of number of affected cases in the USA.
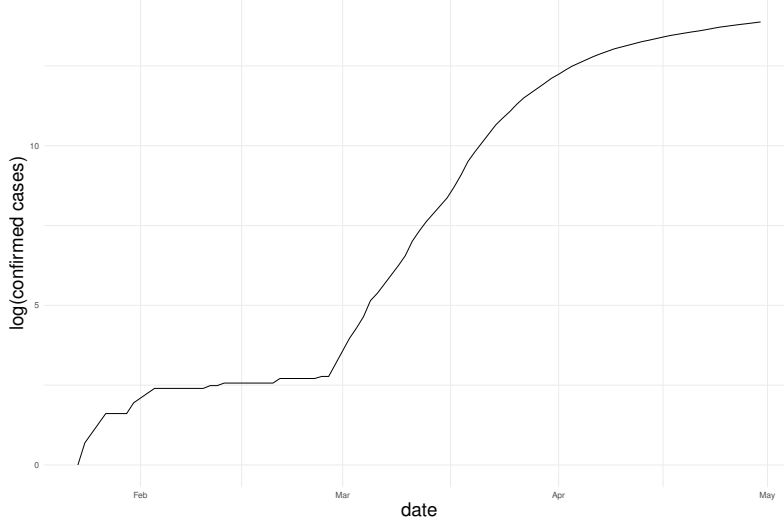
Figure 2: Daily number of COVID-19 affected cases (log-transformed) in the USA.

2.2 Cumulative abnormal return due to the pandemic

In continuation with the previous section, we start with a look at the abnormal returns for the airlines stock prices following the COVID-19 outbreaks. We use event study methodology which is based on the hypothesis that capital markets are efficient in processing information by establishing an appropriate new stock price equilibrium as soon as new information becomes available, cf. Fama et al. (1969). The underlying logic is that the investors in capital markets process the available information on different firm activities and external events, and consider both the impact on current performance and the impact on the future performance of the firms. With more information being available, a stock price index would change and would reflect the investors' revised consensus about future profitability. The changes in investors' beliefs regarding the future profitability of a firm are thus reflected in abnormal returns, which are the risk adjusted returns in excess of the firm's expected return. As pointed out by McWilliams and Siegel (1997), abnormal returns provide a unique way of associating the impact of a particular event on the firm's expected profitability in future periods. Thus, a first step to understand the effect of a pandemic is to measure the abnormal return of the airlines following the outbreak.

Let $R_{s,t}$ denote the return of a specific stock s at time t and $R_{m,t}$ denote the return of the market index at time t. In our analysis, S&P 500 index (data obtained from Yahoo Finance) is used for that purpose. Now, consider the ordinary linear regression:

$$R_{s,t} = \alpha_s + \beta_s R_{m,t} + \epsilon_{s,t}. \tag{2.2}$$

Suppose, T denotes the time of the event (start date of an outbreak) we are interested in. Now, we use all data of the stock s for $t \leq T$ to fit the above model and estimate the coefficients, hereafter denoted by $\hat{\alpha}_s, \hat{\beta}_s$. Then, for all $t' > T$, the abnormal return of the stock s is evaluated by

$$\mathrm{AR}_{s,t'} = R_{s,t'} - \hat{\alpha} - \hat{\beta}_s R_{m,t'}. \tag{2.3}$$

In other words, the abnormal returns are the prediction errors of the model described by eq. (2.2). They denote the returns over and above the return predicted by general market trend on the days past an event of interest. We assume that the abnormal returns are the result of the pandemics and not some other random event occurring on the same day. Further, we compute the cumulative abnormal return (CAR) of stock s corresponding to a pandemic happening at time T. For a window of length w, it is

defined as:

$$\mathrm{CAR}_{s,w} = \sum_{t'=T+1}^{T+w} \left(R_{s,t'} - \hat{\alpha} - \hat{\beta}_s R_{m,t'} \right). \tag{2.4}$$

Figure 3: CAR values for different window-lengths following the four pandemics for the returns of three airlines stocks.

In Figure 3, CARs for different window-lengths (up to 100, which is roughly four months) following the COVID-19 outbreak are presented. For comparison, the same results for 2017-18 flu epidemic and

6

the 2007-08 global recession are included as well. Note that the latter has been the worst financial crisis during the last decade and it affected the airlines industry severely. Wang (2013) can be referred for a relevant in-depth discussion. The graphs portray similar story for all of the three airlines. It is evident that the effect was minimal for 2017-18 flu. On the contrary, COVID-19 has already caused a much more severe abnormal return for all three stocks which is more or less similar to the great recession. It was however unprecedented. One can guess that the extent of public reaction to the pandemic and its affect on the airlines stocks could not be judged properly. That naturally brings forth an important question. Is there another way to develop an appropriate modeling technique which will help us in analyzing and predicting the stock returns? We aim to tackle that issue in the following section.

3 Methods

3.1 Predictor variables

One of the main assumption of this paper is that the public interest and behavior during a pandemic can explain the high level of volatility of the airlines stocks. We quantify this by taking the volume of search history in Google. For that, an appropriate set of keywords is first found out with the help of Twitter data.

Note that Twitter, at present, has more than 300 million active users worldwide, while in the USA, nearly 50 million active users are recorded per month. Over the last few years, many interesting studies have come forward to establish that the Twitter data can be leveraged wisely in financial modeling. For example, Mao et al. (2012) showed that S&P 500 stocks are highly correlated with the daily volume of tweets that mention 'S&P stocks'. A few more papers have established that a sentiment analysis of twitter data can be instrumental in predicting stock market movements. Ranco et al. (2015) and Pagolu et al. (2016) are worth mention in this regard. While these studies make use of public tweets directly in the modeling, our objective is not to assess the public sentiment from the tweets, but to solely get a sense of the main topics of interest for general people. It is true that nowadays Twitter acts as a news source for not only public, but even for mainstream media as well. Refer to Moon and Hadley (2014) for some relevant discussion.

For our purpose, ten official Twitter handles, namely the ones from American Airlines, Delta Airlines, United Airlines, Yahoo Finance, Financial Times, CNBC, CNN, MSNBC, NBC News and Fox News, are selected. These are chosen carefully to reflect three aspects that are of interest to public so far as our study is concerned - details pertaining to traveling, flights and airlines, information regarding economy and stock market, daily news and current affairs. Now, a sample of 30,000 tweets are parsed from the above handles and we do a term-frequency analysis to find out most common terms. At this stage, the stop words, links and tags are removed and the data is cleaned to reflect the main topics of the tweets.

Next, we look at the most frequent words and identify 25 key-phrases that are of public interest for various reasons. From now on, we will call this set of main keywords and phrases as 'topics'. They are listed below. Observe that the topics can be broadly categorized into three categories. First one is the set of topics people would search usually without any pressing concern. We call it 'general' and examples are 'american airlines', 'delta airlines stock' etc. Next, the topics like 'flight status', 'cancellation', 'reservation' etc are used when someone needs quick update or action on travel plans. We would name this category 'travel'. Finally, the topics related to pandemics and relevant concerns, for example 'coronavirus', 'epidemic', 'community spread', are reserved for the category 'COVID'. We are primarily interested in the effect of the third category on the airlines stock prices. The full list of topics are provided below:

- General (hereafter denoted as $\mathcal{G}$): american airlines, delta airlines, united airlines, economy, american airlines stock, delta airlines stock, united airlines stock, stock price, president trump, government

- Travel (hereafter denoted as $\mathcal{T}$): flight cancellation, flight status, flight delay, flight assistance, flight booking, reservation

- COVID (hereafter denoted as $\mathcal{C}$): lockdown, coronavirus, pandemic, community spread, covid, travel ban, shutdown, disease, epidemic

We emphasize that the above categorization is subjective and is done carefully using the tweets shared by the official handles. Now, to do our main analysis, for all of the above topics, we want to obtain the Google search volumes to capture the level of public interest on different days. At this stage, we make a minor adjustment and choose not only these exact phrases, but also five most common queries related to the topics. For example, 'american airlines' is related to the queries 'flight', 'american airlines flight', 'american airlines flights', 'flights' etc. The phrase 'community spread' is related to 'coronavirus', 'coronavirus community spread', 'what is community spread', 'community spread virus' etc. Now, using Google Trends, we compute the daily Google search volume for the 25 topics. Let S_{it} denote the search volume for topic i on tth day. It is worth mention that Google Trends provide the relative interest for every keyword within a topic and for every day on a scale from 1 to 100. So, proper measures were taken to convert the variables to form the daily time series we require. Further, in line with how return series is defined as the change in stock price, we compute daily change in the search volume. We call this 'internet concern' and for topic i on tth day, it is denoted by IC_{it}.

$$IC_{it} = \log S_{it} - \log S_{i,t-1}. \tag{3.1}$$

A much needed exercise at this stage is to ensure that the internet concerns for the topics are in fact useful to predict changes in volatility. We perform Granger causality tests in that regard, cf. Granger (1969). In Table 2, the F-values corresponding to the Granger tests between IC_{it} and daily volatility for the three airlines stocks are reported. Significance at different levels are denoted by asterisks.

We can see that the airlines' names are usually significant, and so are 'united airlines stock', 'stock price' and 'government'. However, other stock price related terms are not significant, suggesting that there has not been any discernible change in daily search volume that can be used to predict shifts in stock return volatility. Similar thing can be said about topics in category $\mathcal{T}$, except for 'flight assistance'. On the other hand, all but two topics in category $\mathcal{C}$ have been found to Granger-cause the stock returns. It is worth noting that the results of the Granger causality tests for the reverse relationship are similar as well.

Now, based on the above findings, we decide to take an appropriately parsimonious approach to avoid too many predictors in the model and aggregate the internet concerns of the significant predictors within every category. At this stage, we consider a topic to be significant if it passes the Granger test for at least one of the three cases. Note that there are only four topics which do not show similar results for all of the airlines and therefore, this simplification will not affect the conclusions much. The three predictors thus created are given below:

$$\mathcal{G}_t = \sum_{i \in \mathcal{G}} b_i IC_{it}, \ \mathcal{T}_t = \sum_{i \in \mathcal{T}} b_i IC_{it}, \ \mathcal{C}_t = \sum_{i \in \mathcal{C}} b_i IC_{it}. \tag{3.2}$$

Here, b_i is a binary term which equal to 1 (or 0) represents whether topic i is significant (or not). We shall also use the daily number of new COVID-19 cases in the USA as a predictor in our model. For day t, this is going to be denoted by $\mathcal{N}_t$.

3.2 Model

We start with a brief discussion of auto-regressive moving average (ARMA) methods which have been used by many researchers to analyze stock returns. Consider a time series $(Y_t)_{1 \leq t \leq n}$ and let us assume it to be stationary. An ARMA model of order (p, q) is defined by

$$Y_t - \phi_1 Y_{t-1} - \ldots - \phi_p Y_{t-p} = e_t + \theta_1 e_{t-1} + \ldots + \theta_{t-q} e_{t-q}. \tag{3.3}$$

Table 2: F-values for the Granger-causality tests between daily volatility of three airlines stocks and internet concern for different topics. Asterisk is used to point out significance at 1% (***), 5% (**) and 10% (*) levels.

Category	Topic	F-value (for AA)	F-value (for DA)	F-value (for UA)
General	american airlines	4.06 ***	3.69 ***	3.2 ***
	delta airlines	2.1 *	1.69	1.32
	united airlines	5.07 ***	4.89 ***	5.65 ***
	economy	0.64	0.65	0.52
	american airlines stock	0.17	0.12	0.34
	delta airlines stock	0.49	0.14	0.34
	united airlines stock	26.17 ***	18.74 ***	28.16 ***
	stock price	1.09	3.63 ***	3.77 ***
	president trump	2.1 *	1.45	1.56
	government	8.82 ***	4.37 ***	5.65 ***
Travel	flight cancellation	0.16	0.06	0.13
	flight status	1.72	0.82	0.92
	flight delay	0.18	0.28	0.36
	flight assistance	3.41 ***	10.18 ***	8.63 ***
	flight booking	0.2	0.9	0.46
	reservation	0.52	0.24	0.19
COVID	lockdown	10.12 ***	3.83 ***	2.99 ***
	coronavirus	80.39 ***	52.12 ***	33.54 ***
	pandemic	5.96 ***	7.02 ***	8.04 ***
	community spread	60.25 ***	71.1 ***	56.93 ***
	covid	42.25 ***	150.68 ***	125.39 ***
	travel ban	0.55	0.29	0.24
	shutdown	3.14 ***	0.68	2.96 ***
	disease	0.49	0.79	0.7
	epidemic	33.5 ***	23.17 ***	24 ***

In the above equation, e_i's are assumed to be zero-mean, independent and identically distributed Gaussian random variables. The ϕ coefficients form the autoregressive (AR) part and the θ coefficients denote the moving average (MA) part. One can also use $\phi(x) = 1 - \phi_1 x - \ldots - \phi_p x^p$ to denote the AR characteristic polynomial, $\theta(x) = 1 + \theta_1 x + \ldots + \theta_q x^q$ to denote the MA characteristic polynomial and B to denote the backward shift operator, in which case, the above model can be written in a simplified form as $\phi(B)Y_t = \theta(B)e_t$. For further reading, (Box et al., 2015, Ch. 3 and 4) is an excellent reference.

One major problem with the ARMA models is that they assume homoskedasticity in the error structure, which is often not the case for financial time series data. To tackle that, Engle (1982) developed the autoregressive conditional heteroskedasticity (ARCH) models to model the time-varying volatility in the data. Later, Bollerslev (1986) extended it to develop generalized auto-regressive conditional heteroskedasticity (GARCH) methods which have been extensively used for volatility forecasting in economics over the last couple of decades. Worthington et al. (2008) and Kathiravan et al. (2019) are two relevant examples of the use of GARCH models in the context of airlines industry. In the current work, we are going to adopt ARMA-GARCH models, where ARMA is used to fit the mean and GARCH to fit the variance. A standard ARMA(p, q)-GARCH(a, b) model is defined mathematically as follows.

$$\phi(B)Y_t = \theta(B)e_t; \ e_t = \sigma_t Z_t; \ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{a} \alpha_i e_{t-i}^2 + \sum_{i=1}^{b} \beta_i \sigma_{t-i}^2. \tag{3.4}$$

Here, Z_t is a white noise process. The coefficients in the model are assumed to be positive. In order to select the most appropriate lags in the mean structure, we would use Akaike information criterion (AIC) defined as $2k - 2\log\hat{L}$, where k is the number of parameters in the model and $\hat{L}$ is the value of the likelihood. The GARCH orders are taken to be $a = 1, b = 1$ throughout the study. Experimentation showed that this specification works the best for forecasting performances. Further, in order to assess the efficacy of the internet concern predictors (see eq. (3.2)), we include their lagged values (or lags up to 4 days) as exogenous regressors in both the mean and the variance structure of the model. Finally, the GARCH$(1,1)$ specification is done through Glosten-Jagannathan-Runkle (GJR) method, introduced by Glosten et al. (1993). Using $\mathbb{I}(\cdot)$ as the indicator function, this model can be written as the following:

$$\phi(B)Y_t = \theta(B)e_t; \ e_t = \sigma_t Z_t; \ \sigma_t^2 = \omega + (\alpha + \gamma\mathbb{I}(e_{t-1} < 0)e_{t-1}^2 + \beta\sigma_{t-1}^2. \tag{3.5}$$

To summarize, we start with a unit root test to check for stationarity and in case of nonstationarity, an initial differencing step can be done. Then, we use an ARMA-GARCH model for the return series where the mean structure is defined with lagged values of $\mathcal{G}_t, \mathcal{T}_t, \mathcal{C}_t$ as exogenous predictors and the ARMA orders are chosen based on AIC, while the variance structure is GARCH with order $(1,1)$ and is defined with same exogenous predictors. Hereafter, for convenience, this model will be called MICP (model with internet concern predictors).

For comparison purposes, we also use two other types of ARMA-GARCH models. One of them is an ARMA-GARCH model with lagged values of daily COVID-19 cases ($\mathcal{N}_t$) as exogenous regressors in both mean and variance structures. We call it MCASES. Similarly, another ARMA-GARCH model where only the COVID internet concern predictor ($\mathcal{C}_t$) is taken as the exogenous regressor is used as well. This is denoted by MCOVID. Finally, a standard ARMA-GARCH model without any exogenous predictor is used as Baseline model in our analysis as well.

3.3 Forecast evaluation using model confidence set

In this section, we discuss the model confidence sets (MCS) procedure (Hansen et al. (2011)) which is used to test if one of the candidate models has superior predictive power than others.

Let M denote the set of candidate models we consider. The forecasting efficiency of these models are evaluated under some loss function for the test set P. Let $L_{m,t}$ denote the loss associated with $m \in M, t \in P$. The relative performance variables are then defined as $d_{m_1 m_2, t} = L_{m_1, t} - L_{m_2, t}$. Let $\mathbb{E}(d_{m_1 m_2, t})$ be the expectation of $d_{m_1 m_2, t}$ under the loss function. Then, the objective of the MCS procedure is to choose a model set M^* such that

$$M^* = \left\{ m \in M : \mathbb{E}(d_{mm', t}) \leq 0, \forall \ m' \in M \right\}. \tag{3.6}$$

The procedure follows a sequence of significance tests with null hypothesis of the form

$$H_{0, M'} : \mathbb{E}(d_{mm', t}) = 0 \ \ \forall \ m, m' \in M' \subset M. \tag{3.7}$$

At every step, an equivalence test $\delta_{M'}$ is conducted to test if any two models in M' perform equally well under the loss function. If $\delta_{M'}$ is rejected, an elimination rule $e_{M'}$ eliminates the model with poor performance. After repeating the tests as needed, the set of surviving models M^* is obtained. Throughout this paper, we use a significance level of 0.1 for the MCS procedure. So far as the test statistic is concerned, the authors suggested different choices and we are going to use the range statistic T_R. If $\bar{d}_{mm'}$ and $v(d_{mm'})$ are respectively the mean and variance of $d_{mm', t}$ for $t \in P$, then

$$T_R = \max_{m, m' \in M'} \frac{|\bar{d}_{mm'}|}{\sqrt{v(d_{mm'})}}. \tag{3.8}$$

Note that the asymptotic distribution of T_R is nonstandard, and it depends on some nuisance parameters. However, as the authors pointed out, it can be estimated using bootstrap methods that implicitly

solve the nuisance parameter problem. For implementation of the above procedure using bootstrap methods, the R package MCS by Bernardi (2017) is used.

As a last piece of the MCS procedure, we need to choose appropriate loss functions. Earlier, it has been discussed (see Diebold and Mariano (2002)) why a single loss function is not the best way to distinguish advantages and disadvantages of the predictive power for various models. In light of that, we use three different loss functions in our work. They are mean squared error (MSE), mean absolute error (MAE) and quasi-likelihood (QLIKE). Let y_t be the daily volatility of a stock price and $\hat{y}_t$ be the corresponding forecast, for $t \in P$, the test set. Then, the above loss functions are defined as follows:

$$\text{MSE} = \frac{1}{|P|} \sum_{t \in P} (\hat{y}_t - y_t)^2. \tag{3.9}$$

$$\text{MAE} = \frac{1}{|P|} \sum_{t \in P} |\hat{y}_t - y_t|. \tag{3.10}$$

$$\text{QLIKE} = \frac{1}{|P|} \sum_{t \in P} \left(\log \hat{y}_t + \frac{y_t}{\hat{y}_t} \right). \tag{3.11}$$

4   Results

Recall that all data before 12th April, 2020 are used for training purposes and the rest are kept as test set. We use four candidate models as mentioned in Section 3.2. For all of these models, GARCH order is taken as $(1,1)$ whereas the ARMA order is chosen by the least value of AIC. In Table 3, the specifications and the in-sample AIC of these four types of candidate models are presented.

Table 3: List of ARMA-GARCH models used to analyze and predict the returns and volatility of three airlines stocks. Predictors are used in both mean and variance structure.

Data	Model	Mean model	Variance model	Predictors	AIC
AA	Baseline	ARMA$(4,2)$	GARCH$(1,1)$		4.54
	MCASES	ARMA$(4,2)$	GARCH$(1,1)$	$\mathcal{N}_t$	4.48
	MCOVID	ARMA$(4,2)$	GARCH$(1,1)$	$\mathcal{C}_t$	4.48
	MICP	ARMA$(4,2)$	GARCH$(1,1)$	$\mathcal{G}_t, \mathcal{T}_t, \mathcal{C}_t$	4.49
DA	Baseline	ARMA$(3,2)$	GARCH$(1,1)$		4.02
	MCASES	ARMA$(3,4)$	GARCH$(1,1)$	$\mathcal{N}_t$	3.94
	MCOVID	ARMA$(4,4)$	GARCH$(1,1)$	$\mathcal{C}_t$	3.95
	MICP	ARMA$(4,3)$	GARCH$(1,1)$	$\mathcal{G}_t, \mathcal{T}_t, \mathcal{C}_t$	3.96
UA	Baseline	ARMA$(4,2)$	GARCH$(1,1)$		4.37
	MCASES	ARMA$(3,2)$	GARCH$(1,1)$	$\mathcal{N}_t$	4.21
	MCOVID	ARMA$(3,2)$	GARCH$(1,1)$	$\mathcal{C}_t$	4.21
	MICP	ARMA$(3,2)$	GARCH$(1,1)$	$\mathcal{G}_t, \mathcal{T}_t, \mathcal{C}_t$	4.22

Maximum order of 4 is used while choosing the best ARMA orders. It has been found that for a fixed GARCH order and regressors, the AIC values for different ARMA specifications are indeed very close. Hence, it should not have significant role in the forecasting performance. Also, the AIC values decrease only marginally after introducing the predictors. The conditional volatility of the stock returns in the training set, as modeled by the four candidate models, are presented in Figure 4 (for AA), Figure 5 (for DA), and Figure 6 (for UA). It is clear that all models capture the peaks and troughs in the stock price movement in a comparable way. Especially towards the end of the training period, the high level of volatility can be explained by the models as well. Detailed results of the estimated coefficients and standard errors of all these models are provided in the Appendix.
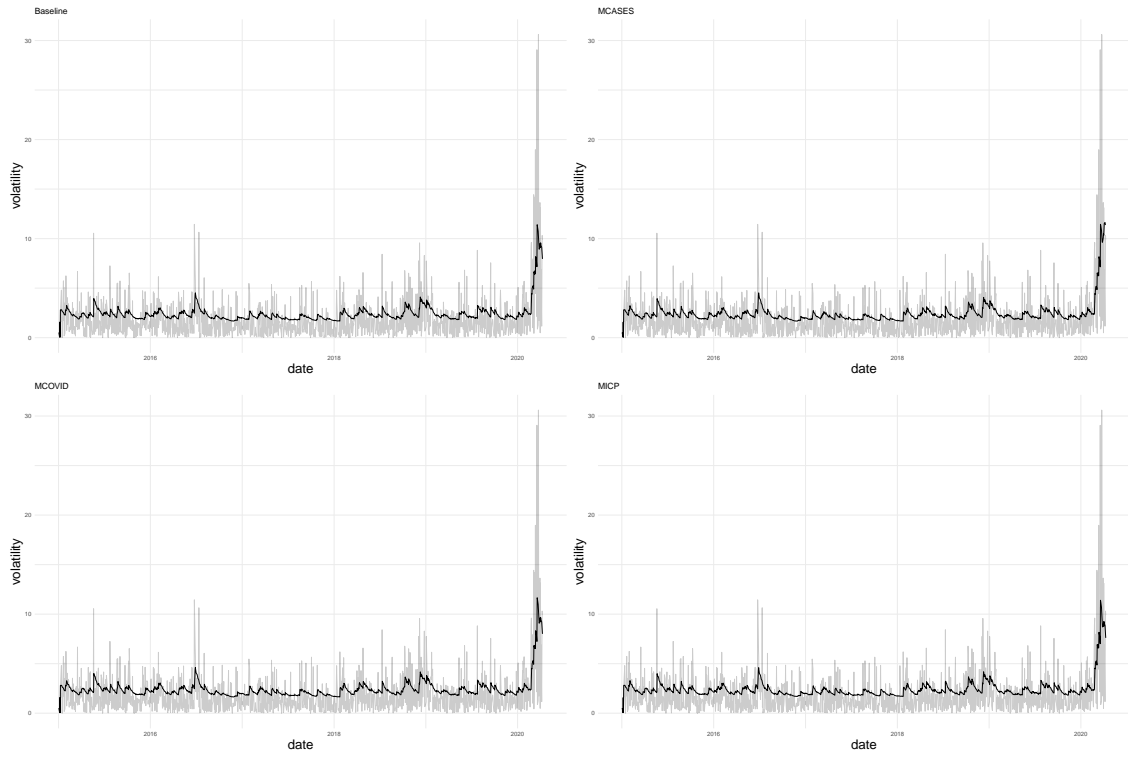
Figure 4: Conditional volatility for AA stock returns, as modeled by four candidate models.
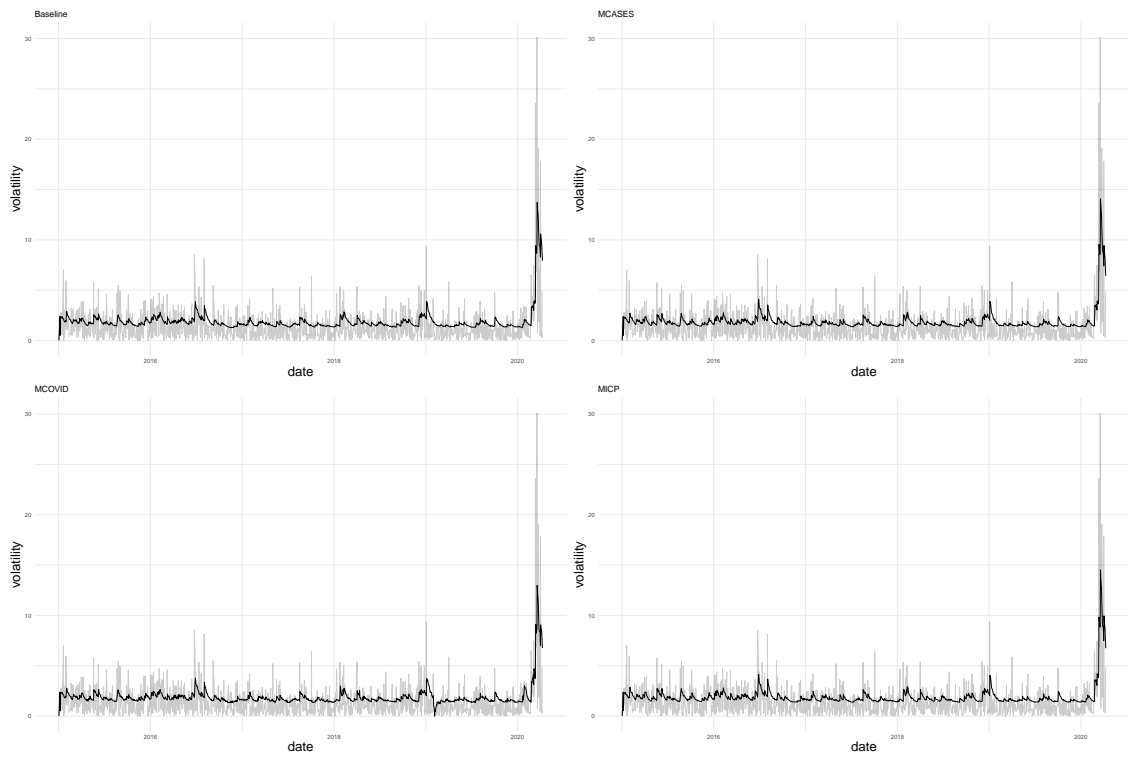

Figure 5: Conditional volatility for DA stock returns, as modeled by four candidate models.
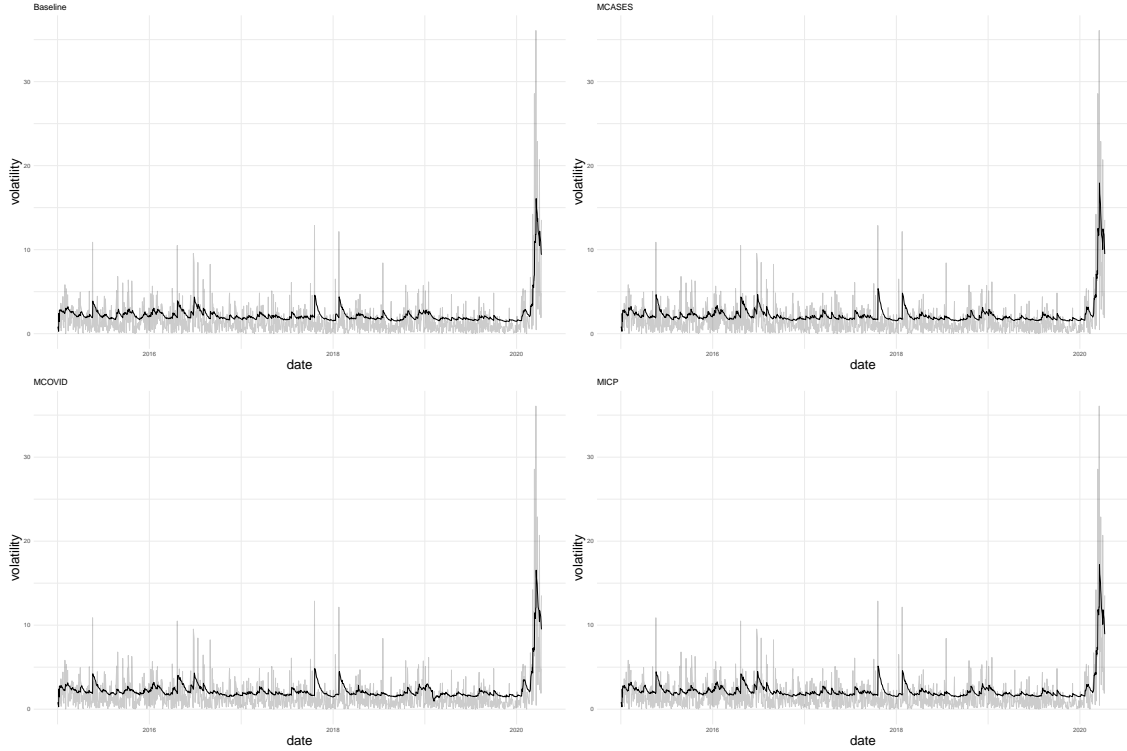
Figure 6: Conditional volatility for UA stock returns, as modeled by four candidate models.

We now turn our attention to forecasting performance, the more important aspect of the analysis. First, the results of the MCS procedure are presented in Table 4. For every model, we get the predicted values of the volatility and compare it to the observed ones from the test set. Three different loss functions and a 10% significance level, as mentioned in the last section, are used in these tests. Thus, p-values of greater than 0.1 indicate which models survive the MCS test under the specific loss function and the test statistics T_R, while a p-value of 1 would indicate a model performing the best out of all candidates.

Table 4: Results of the model confidence set (MCS) test for the candidate models, according to three different loss functions. Best models are marked in bold.

Data	Model	For MSE loss	For MAE loss	For QLIKE loss
AA	Baseline	0.00	0.00	0.00
	MCASES	0.00	0.00	0.00
	MCOVID	0.00	0.00	0.00
	MICP	**1.00**	**1.00**	**1.00**
DA	Baseline	0.30	0.24	0.66
	MCASES	**1.00**	**1.00**	**1.00**
	MCOVID	**1.00**	**1.00**	**1.00**
	MICP	**1.00**	**1.00**	**1.00**
UA	Baseline	0.00	0.00	0.00
	MCASES	0.00	0.00	0.00
	MCOVID	0.00	0.00	0.00
	MICP	**1.00**	**1.00**	**1.00**

It can be observed that for all cases, MICP survives the MCS test. In fact, it performs the best of

the four in all cases. The baseline model, except for DA stock returns, does not survive the MCS test. Same thing is true for MCOVID and MCASES as well. These two however, unlike the baseline model, perform equally good as MICP for DA stock returns. These results are very straight-forward to establish that MICP is the single best model across all loss functions.

Next, we take a look at the improvement in the three loss functions for volatility forecasting from the baseline model for the three stock returns. Table 5 lists these values.

Table 5: Percentage improvement for the three candidate models from the baseline model under three different loss functions.

Data	Model	MSE	MAE	QLIKE
AA	MCASES	−443.79%	−144.72%	−12.29%
	MCOVID	−0.78%	−0.32%	−0.05%
	MICP	6.55%	3.70%	0.43%
DA	MCASES	23.65%	15.36%	1.65%
	MCOVID	21.70%	13.51%	1.37%
	MICP	21.98%	13.32%	1.57%
UA	MCASES	13.06%	7.93%	0.86%
	MCOVID	4.37%	1.73%	0.28%
	MICP	14.72%	9.21%	0.97%

For AA, only MICP performs better than the baseline model in terms of all three loss functions. For the other two airlines stock returns, all three models beat the baseline model. MCASES does terribly for all loss functions in case of AA. For DA stock returns though, it has much better improvement over other two models in terms of MAE and MSE whereas under QLIKE loss, the three models see comparable improvement. Finally, for UA data, MICP beats MCASES and MCOVID ubiquitously, but MCASES does better than MCOVID. Further, note that for DA stock returns, the improvement in MSE and MAE are much more substantial than AA or UA stock returns. On the other hand, a closer look at the results has shown that MCASES, which uses daily number of COVID-19 cases as the predictor, always tends to predict high volatility, which is not the reality for AA stock returns, and hence this model fails miserably there.

Overall, we see that the proposed model outperforms the other candidate models. The actual number of daily COVID-19 cases is not a good predictor for volatility forecasting for airlines stock prices as it suffers heavily from overforecasting issues. With the number of cases growing exponentially in the USA, a much better strategy is to quantify the effect of the pandemic on general public by using the internet concern variable defined in this paper. Moreover, based on the above findings, one can say that the internet concern for different aspects, for example general, travel related and pandemic related, contain more useful information than the internet concern for only pandemic related aspect.

5   Conclusion

This paper serves two main purposes. On one hand, it is an endeavor to understand the severity of COVID-19 on airlines stock prices while on the other, the proposed model provides a new way to extract relevant information from people's behavior on the internet to construct useful predictors for forecasting stock return volatility. It is found that the information obtained from the internet search volume on certain topics provide higher predictive accuracy. Thus, this study can provide a useful analytics tool for market participants, policymakers, and market regulators to understand the movement of airlines stock market in the future. In fact, not only for a global issue like COVID-19 pandemic, but this type of approach can also be used to understand and predict the stock market from other aspects, such as political activities, terrorism, crashes or accidents, in short whatever can be thought to leave a significant

effect on the airlines stock prices.

An attractive feature of the proposed method is its flexibility in incorporating more information. For instance, we aggregated the topics within a category to get a holistic view of the internet concern in the three broad aspects. Instead, one can choose to treat the IC_{it} series separately for every topic, and use those data in the model. Adding more topics within a category or considering other categories is also possible to gather more information in the internet concern predictors. On that note, a limitation of the proposed method is that the Twitter data analysis has been done in a limited capacity. We collected approximately thirty thousand tweets in this analysis whereas more than millions of tweets are available to conduct a more extensive study. Although it demands a lot more computational burden, a more detailed analysis of Twitter data can in fact be used to formulate a time-dependent set of topics, which can be subsequently used to create an insightful predictor that captures the internet concern on trending topics more appropriately. This has the potential to be used efficiently in short term volatility forecasting, and is a good future direction to this work. Another interesting extension of the current framework is to use GARCH MIDAS model, which has been used in the context of volatility forecasting in recent times, in conjunction with the internet search predictors.

Finally, coming to the main data we analyzed in this paper, one can see that there has been a much higher level of volatility in the stock price index of three of the biggest airlines companies from the USA, and that can be explained by the internet concern on the relevant topics related to the pandemic. With COVID-19 cases increasing everyday, there is a chance that in near future, the governments will have to impose new policies to contain the virus. That is bound to cause problems for the aviation industry and the major companies will incur losses once again. It can be hypothesized that the other big companies are also going to suffer similarly while the smaller airlines will take a greater hit should that happen. Earlier in March 2020, British regional airlines Flybe ran out of cash and entered administration. In the USA, two small-scale services Compass Airlines and Trans States Airlines suffered the same fate in April. Elsewhere in the world, Virgin Australia is another company who went into bankruptcy as an effect of the coronavirus carnage. Many other similar airlines companies are currently hanging by the thread and are hoping to see the pandemic to come to an end.

Declaration of interest statement

No potential conflict of interest was reported by the authors.

Data availability statement

The data that support the findings of this study are openly available in JHU-CSSE (2020) and Yahoo Finance at https://in.finance.yahoo.com/.

References

Scott Baker, Nicholas Bloom, Steven J Davis, Kyle Kost, Marco Sammon, and Tasaneeya Viratyosin. The unprecedented stock market reaction to covid-19. *Covid Economics: Vetted and Real-Time Papers*, 1 (3), 2020.

Leopoldo Catania & Mauro Bernardi. *MCS: Model Confidence Set Procedure*, 2017. URL https://CRAN.R-project.org/package=MCS. R package version 0.1.3.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31 (3):307–327, 1986.

George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

Matteo Chinazzi, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science*, 2020.

Soudeep Deb and Manidipa Majumdar. A time series method to analyze incidence pattern and estimate reproduction number of covid-19. *arXiv preprint arXiv:2003.10655*, 2020.

Francis X Diebold and Robert S Mariano. Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134–144, 2002.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

Martin Enserink and Kai Kupferschmidt. With covid-19, modeling takes on life and death importance, 2020.

Eugene F Fama, Lawrence Fisher, Michael C Jensen, and Richard Roll. The adjustment of stock prices to new information. *International economic review*, 10(1):1–21, 1969.

Nuno Fernandes. Economic effects of coronavirus outbreak (covid-19) on the world economy. *Available at SSRN 3557504*, 2020.

Lawrence R Glosten, Ravi Jagannathan, and David E Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993.

Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438, 1969.

Veronica Guerrieri, Guido Lorenzoni, Ludwig Straub, and Iván Werning. Macroeconomic implications of covid-19: Can negative supply shocks cause demand shortages? Technical report, National Bureau of Economic Research, 2020.

Peter R Hansen, Asger Lunde, and James M Nason. The model confidence set. *Econometrica*, 79(2): 453–497, 2011.

Stefano Maria Iacus, Fabrizio Natale, Carlos Santamaria, Spyridon Spyratos, and Michele Vespe. Estimating and projecting air passenger traffic during the covid-19 coronavirus outbreak and its socio-economic impact. *Safety Science*, page 104791, 2020.

[Dataset] JHU-CSSE. 2019 novel coronavirus covid-19 (2019-ncov) data repository by johns hopkins csse, 2020. URL https://github.com/CSSEGISandData/COVID-19. Last accessed on 1st May, 2020.

Chinnadurai Kathiravan, Murugesan Selvam, Balasundram Maniam, Sankaran Venkateswar, et al. Relationship between crude oil price changes and airlines stock price: The case of indian aviation industry. *International Journal of Energy Economics and Policy*, 9(5):7–13, 2019.

Qianying Lin, Shi Zhao, Daozhou Gao, Yijun Lou, Shu Yang, Salihu S Musa, Maggie H Wang, Yongli Cai, Weiming Wang, Lin Yang, et al. A conceptual model for the outbreak of coronavirus disease 2019 (covid-19) in wuhan, china with individual reaction and governmental action. *International Journal of Infectious Diseases*, 2020.

Yuexin Mao, Wei Wei, Bing Wang, and Benyuan Liu. Correlating s&p 500 stocks with twitter data. In *Proceedings of the first ACM international workshop on hot topics on interdisciplinary social networks research*, pages 69–72, 2012.

Warwick J McKibbin and Roshen Fernando. The global macroeconomic impacts of covid-19: Seven scenarios. *Available at SSRN 3557504*, 2020.

Abagail McWilliams and Donald Siegel. Event studies in management research: Theoretical and empirical issues. *Academy of management journal*, 40(3):626–657, 1997.

Soo Jung Moon and Patrick Hadley. Routinizing a new technology in the newsroom: Twitter as a news source in mainstream media. *Journal of Broadcasting & Electronic Media*, 58(2):289–305, 2014.

Venkata Sasank Pagolu, Kamal Nayan Reddy, Ganapati Panda, and Babita Majhi. Sentiment analysis of twitter data for predicting stock market movements. In *2016 international conference on signal processing, communication, power and embedded system (SCOPES)*, pages 1345–1350. IEEE, 2016.

Gabriele Ranco, Darko Aleksovski, Guido Caldarelli, Miha Grčar, and Igor Mozetič. The effects of twitter sentiment on stock price returns. *PloS one*, 10(9), 2015.

Arkaprava Roy and Sayar Karmakar. Bayesian semiparametric time varying model for count data to study the spread of the covid-19 cases. *arXiv preprint arXiv:2004.02281*, 2020.

Muhammad Adnan Shereen, Suliman Khan, Abeer Kazmi, Nadia Bashir, and Rabeea Siddique. Covid-19 infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*, 2020.

Catrin Sohrabi, Zaid Alsafi, Niamh O'Neill, Mehdi Khan, Ahmed Kerwan, Ahmed Al-Jabir, Christos Iosifidis, and Riaz Agha. World health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19). *International Journal of Surgery*, 2020.

Yu Shan Wang. The impact of crisis events on the stock returns volatility of international airlines. *The Service Industries Journal*, 33(12):1206–1217, 2013.

Andrew C Worthington et al. The impact of natural events and disasters on the australian stock market: A garch-m analysis of storms, floods, cyclones, earthquakes and bushfires. *Global Business and Economics Review*, 10(1):1, 2008.

Joseph T Wu, Kathy Leung, and Gabriel M Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet*, 395(10225):689–697, 2020.

Lean Yu, Yaqing Zhao, Ling Tang, and Zebin Yang. Online big data-driven oil consumption forecasting with google trends. *International Journal of Forecasting*, 35(1):213–223, 2019.

Appendix

Coefficient estimates from fitted models

Table A1: Estimates of the coefficients in the four models for AA data. Standard errors are provided in parentheses. $*$ denotes significant effect at 5% level of significance.

Variable	Baseline	MCASES	MCOVID	MICP
intercept	$-0.06(0.00)^*$	$0.00(0.01)$	$-0.02(0.04)$	$-0.03(0.04)$
AR1	$-1.84(0.00)^*$	$-1.85(0.00)^*$	$-1.85(0.00)^*$	$-1.86(0.00)^*$
AR2	$-0.92(0.00)^*$	$-0.93(0.00)^*$	$-0.93(0.00)^*$	$-0.95(0.00)^*$
AR3	$-0.03(0.00)^*$	$-0.03(0.00)^*$	$-0.02(0.00)^*$	$-0.04(0.00)^*$
AR4	$-0.04(0.00)^*$	$-0.04(0.00)^*$	$-0.03(0.00)^*$	$-0.04(0.00)^*$
MA1	$1.92(0.00)^*$	$1.92(0.00)^*$	$1.92(0.00)^*$	$1.92(0.00)^*$
MA2	$1.01(0.00)^*$	$1.01(0.00)^*$	$1.01(0.00)^*$	$1.01(0.00)^*$
$\mathcal{N}_t$ (mean)		$0.00(0.00)$		
$\mathcal{C}_t$ (mean)			$-0.04(0.18)$	$0.00(0.02)$
$\mathcal{T}_t$ (mean)				$-0.02(0.12)$
$\mathcal{G}_t$ (mean)				$-0.09(0.21)$
$\mathcal{N}_t$ (variance)		$0.00(0.00)$		
$\mathcal{C}_t$ (variance)			$0.00(4.21)$	$0.00(0.21)$
$\mathcal{T}_t$ (variance)				$0.00(0.93)$
$\mathcal{G}_t$ (variance)				$0.00(2.09)$
ω	$0.23(0.66)$	$0.25(2.11)$	$0.22(1.36)$	$0.27(0.97)$
α	$0.00(0.27)$	$0.00(0.51)$	$0.00(0.41)$	$0.00(0.19)$
β	$0.91(0.35)$	$0.91(0.84)$	$0.91(0.67)$	$0.90(0.38)$
γ	$0.10(0.04)$	$0.09(0.06)$	$0.10(0.21)$	$0.10(0.10)$
shape		$4.95(0.62)^*$	$5.70(3.63)$	$5.56(1.64)^*$

Table A2: Estimates of the coefficients in the four models for DA data. Standard errors are provided in parentheses. $*$ denotes significant effect at 5% level of significance.

Variable	Baseline	MCASES	MCOVID	MICP
intercept	$-0.01(0.06)$	$-0.03(0.00)^*$	$0.05(0.00)^*$	$-0.03(0.00)^*$
AR1	$1.75(0.02)^*$	$-0.41(0.00)^*$	$0.24(0.01)^*$	$-0.41(0.00)^*$
AR2	$-1.02(0.01)^*$	$0.40(0.00)^*$	$0.39(0.00)^*$	$0.40(0.00)^*$
AR3	$0.02(0.00)^*$	$0.99(0.00)^*$	$0.36(0.01)^*$	$0.98(0.00)^*$
AR4			$-0.90(0.04)^*$	$0.00(0.01)^*$
MA1	$-1.75(0.01)^*$	$0.39(0.00)^*$	$-0.25(0.00)^*$	$0.40(0.00)^*$
MA2	$1.01(0.02)^*$	$-0.43(0.00)^*$	$-0.41(0.01)^*$	$-0.42(0.00)^*$
MA3		$-1.01(0.00)^*$	$-0.34(0.01)^*$	$-1.01(0.00)^*$
MA4		$0.02(0.00)^*$	$0.93(0.00)^*$	
$\mathcal{N}_t$ (mean)		$0.00(0.00)$		
$\mathcal{C}_t$ (mean)			$-0.09(0.00)^*$	$-0.03(0.08)$
$\mathcal{T}_t$ (mean)				$-0.03(0.04)$
$\mathcal{G}_t$ (mean)				$0.01(0.08)$
$\mathcal{N}_t$ (variance)		$0.00(0.00)$		
$\mathcal{C}_t$ (variance)			$0.42(0.01)^*$	$0.00(0.66)$
$\mathcal{T}_t$ (variance)				$0.00(0.52)$
$\mathcal{G}_t$ (variance)				$0.00(0.13)$
ω	$0.20(1.87)$	$0.27(0.01)^*$	$0.22(1.36)$	$0.30(0.24)$
α	$0.00(1.40)$	$0.00(0.00)^*$	$0.00(0.41)$	$0.01(0.02)$
β	$0.88(2.03)$	$0.84(0.19)^*$	$0.85(0.11)^*$	$0.83(0.05)^*$
γ	$0.13(0.77)$	$0.13(0.17)$	$0.12(0.00)^*$	$0.13(0.19)$
shape		$4.66(1.04)^*$	$4.82(4.68)$	$4.63(0.65)^*$

Table A3: Estimates of the coefficients in the four models for UA data. Standard errors are provided in parentheses. ∗ denotes significant effect at 5% level of significance.

Variable	Baseline	MCASES	MCOVID	MICP
intercept	$-0.10(0.00)^*$	$-0.01(0.00)^*$	$0.01(0.00)^*$	$0.03(0.48)$
AR1	$0.10(0.00)^*$	$1.98(0.04)^*$	$1.98(0.04)^*$	$1.98(0.05)^*$
AR2	$0.95(0.00)^*$	$-1.09(0.02)^*$	$-1.09(0.02)^*$	$-1.09(0.03)^*$
AR3	$-0.10(0.00)^*$	$0.06(0.00)^*$	$0.06(0.00)^*$	$0.06(0.00)^*$
AR4	$0.05(0.00)^*$			
MA1	$0.00(0.00)^*$	$-1.96(0.01)^*$	$-1.96(0.03)^*$	$-1.96(0.01)^*$
MA2	$-1.01(0.00)^*$	$1.01(0.01)^*$	$1.01(0.01)^*$	$1.01(0.02)^*$
$\mathcal{N}_t$ (mean)		$0.00(0.00)$		
$\mathcal{C}_t$ (mean)			$-0.11(0.00)^*$	$-0.07(5.04)$
$\mathcal{T}_t$ (mean)				$-0.04(1.09)$
$\mathcal{G}_t$ (mean)				$0.01(0.50)$
$\mathcal{N}_t$ (variance)		$0.00(0.00)$		
$\mathcal{C}_t$ (variance)			$0.35(1.11)$	$0.00(23.55)$
$\mathcal{T}_t$ (variance)				$0.00(4.62)$
$\mathcal{G}_t$ (variance)				$0.00(0.44)$
ω	$0.23(0.63)$	$0.26(9.93)$	$0.20(1.82)$	$0.24(0.77)$
α	$0.07(0.09)$	$0.02(1.35)$	$0.01(0.62)$	$0.02(2.90)$
β	$0.87(0.22)^*$	$0.87(4.50)$	$0.89(1.40)$	$0.87(8.03)$
γ	$0.04(0.23)$	$0.13(2.94)$	$0.11(1.19)$	$0.12(5.10)$
shape		$4.47(3.34)$	$4.46(1.27)$	$4.51(16.86)$