Sentiment Analysis of COVID-19 Pandemic on the Stock Market

Prajwal Eachempati\textsuperscript{a}, Praveen Ranjan Srivastava\textsuperscript{a}, and Prabin Kumar Panigrahi\textsuperscript{b}

https://doi.org/10.37625/abr.24.1.141-165

ABSTRACT

COVID-19 is a dreadful infectious disease, morphed into an economic crisis causing extensive and longstanding ramifications across global markets. Investors continue to hear about COVID-19 and its impact in one corner of the globe or the other for a long time. Though the effects of COVID-19 started in December 2019 in Wuhan, China, global markets did not respond actively till W.H.O officially declared on March 11, 2020, that the COVID-19 outbreak is a global pandemic. These multi-channel events have eroded investor sentiment, tanking the global stock markets. This article uses a machine learning approach to Twitter to analyze and follow investor sentiment that has guided the market to the new low during the first 150 days of the COVID-19 era. The only respite for recovery of financial markets is the lowering of COVID-19 infected cases for the time being till a vaccine is developed for the virus.

KEYWORDS

COVID, Stock Market, Social Media, Machine Learning, Natural Language Processing

INTRODUCTION

Successful investing is all about managing risk by increasing risk-taking ability. Unlike other stock market risk events that impact a particular exchange or industry segment or firm for a short period, COVID-19 carried a global systemic risk. Economic slowdown due to COVID19 resulted in cash flow slump and spiked default rates across various business segments, creating multiple 'epicentres' of market shocks. Amid panic actions of the investors, some global players who are contrarians took the opportunity to increase their holdings in some valuable stocks across exchanges. The markets have to put up with the bad news as long as the microscopic malice of the pandemic is not mitigated by the collective efforts of the global population to guard themselves and medical research to find the cure. Governments of various countries play a critical role in elevating investors’ spirits and for alleviating their fears. Thus, COVID19 is viewed as different from dotcom corrections in 1999-2000, the global financial crisis of 2008 or energy price correction or Ebola during 2014.

INVESTOR SENTIMENT

The Investor sentiment is critical for timing the investment decisions and is an expression of irrational expectations of a stock’s risk-return profile that is not justified by available information. Do investor’s fears due to COVID-19 are irrational that caused markets to oscillate and collapse? How far can investor sentiment impact markets during bubbles and collapse of the market? Has the COVID-19 fever lead to
a market correction and impacted only the irrationality in the market concerning some costly stocks? Investors, in general, find it difficult to predict market extreme positions. Hence, sentiment cannot capture market turbulence or volatility in extreme positions of high or low.

**OBJECTIVES OF THE STUDY**

The study is conducted to analyze the varying Twitter sentiment (Kharde and Sonawane, 2016) on the Coronavirus during pre-lockdown and lockdown period impacting the sentiment on stock market for both the periods. We validate the sentiment impact on the interrelation of various markets to find out if the past global leadership pattern would continue to influence the financial markets after the COVID-19 phase.

The paper is structured as follows:

The literature review detailing the existing studies the impact on stock markets due to pandemics both in the past and currently due to COVID-19 in the study. The data collection and research methodology are elucidated. The results and discussion are illustrated. The study is then concluded and the scope for future research is discussed. This is followed by the references.

**LITERATURE REVIEW**

Behavioral finance literature has shown many sub-optimal decision criteria. Barber and Odean (2005) have stated that investors trade stocks in the limelight because attention is a scarce resource. Record of past returns and high turnover strongly influence trading activity. Instead, what is in the limelight today is the bad news relating to the effects of the pandemic and past performance of stocks has no significant bearing on current crisis. Jiang, Lee, Martin & Zhou (2019) have evaluated the impact of the investment manager’s sentiment on market returns and it is a fascinating study to periodically evaluate their sentiment (Schadewitz et al., 2002) using psycho-traumatic tests during COVID era. The literature on the impact of COVID-19 on financial markets is double edged. The studies describe how the market reacted in similar pandemic incidents in the past and attempt to describe the possible impact of the pandemic on various industry segments. According to Kwon et al., (2020), market experiences during COVID-19 will "shape future investigations of tail risk and financial markets" if similar pandemic outbreaks occur in the future. Due to market action, bad news on COVID-19 has ramifications on all exchanges. The textual information that impacts investor sentiment directly influences the liquidity and volume of trades. Though Chung, Hung & Yeh (2012) have concluded that in a’ recession state, the predictive power of sentiment is generally insignificant’, we need to test these findings in the context of COVID-19.

Though theories on sentiment suggest that it is not uniform among all classes of investors and is heterogeneous, COVID-19 proved otherwise. Though retail traders tend to follow institutional strategists, they may not be able to take advantage of the sentiment—the delayed market reaction to COVID-19(December 19, 2019 to March 11, 2020) supports this view. Edelen, Marcus, and Tehranian (2010) found that retail investor sentiment does not consider a stock’s fundamentals and causes “Noise” (Black, 1986). As investors follow recommendations of investment newsletters, and other reports of market professionals, they are a proxy for institutional investor sentiment. Contrary views are found in Clarke and Statman (1998) and Brooks (2008). Even in such cases, COVID-19 is an exception where news authors were in unison about the impact of the pandemic, which culminated in the collapse of the market.

In Baker et al., (2020), the authors used text-based methods to conclude that COVID-19 has devastated the stock markets due to policy response to the pandemic outbreak. The authors considered the available data from 1900 with a focus on data from October 1987 and December 2008,
1929 to 1930s, to arrive at the above conclusions. On examining contemporary newspaper accounts, the paper states that significant daily changes in the market to pandemic-related developments were not highlighted in earlier periods as during COVID-19. The severity of the pandemic, speedier diffusion of information, the interconnectedness of the global economy, investor behavior and policy response, travel restrictions, social distancing, and other containment norms have caused the markets to tank.

Estimates in the pre-COVID-19 era, show that the global Algorithmic Trading (AT) market might grow significantly between 2018 ($11.1 billion) and 2026 at a CAGR of 11.1% because cloud-based services for algorithmic trading would be available in the markets. In Mukerji, Chung, Walsh & Xiong (2019) and Bazeley and Jackson (2013), the implications of Algorithmic trading on sentiment-analysis of AT behavior with AT trading strategies during specific market conditions and measure the impact of AT on market quality were considered. They observed that as the share of AT in their simulation model reaches 10% of market trades, initially, liquidity increases. But if the share of trades using AT crosses 10%, liquidity increases marginally. Further, statistical arbitrage appears to guide the trade outcomes without anchoring to fundamentals. Thus, at the estimated current level of AT globally at 46% of trades, (World Economic Forum report) noise trading behavior and its impact appear unavoidable as traders tend to ignore Fundamental strengths of stock and gravitated to market sentiment and socio-economic events like COVID-19. The market microstructure is responsible partly for price discovery and formation. As per a report, Agency algorithmic traders (AAT) are better estimators of the current stock price than proprietary algorithmic traders (PAT), suggesting that the institutional investors control stock-specific information. Primarily AAT incorporates the buy-sell transactions at an institutional level and PAT help in stabilizing the market liquidity based on current demand-supply in the market (Nawn & Banerjee, 2017; Shrestha, Subramaniam & Thiyagarajan, 2020). The findings constitute an important element of the sentiment formation process. We consider the impact of COVID-19 from this perspective in the sentiment analysis model developed in the article.

**THE GREAT CRASH OF 1929 AND OTHER RETROSPECTIVE NEWS**

In Hull and White (1990), the authors investigated whether fundamentals or the bubble drove the bull market upwards to reach a market crash point causing boom and bust during great crash of 1929. The rapid economic growth of the 1920s and mania to invest in the bull market led to bubble formation. 'Bubble' in the market caused the crash. The technological and structural changes in the industry in the 1920s developed investors' hope about higher earnings and dividend payouts, setting the stage for a boom. But valuing fundamentals is difficult, which led to the bubble. Instead of allowing the bubble to burst on its own accord, policies of the Federal Reserve helped to push the economy into recession. Thus, the lessons from the past indicate that Government policies and intervention during COVID-19 have a critical role in injecting confidence in the market traders. The crash in 2020 is entirely for different reasons, and investors’ irrationality did not cause the crash. The market computed the consequences of the COVID-19 and expressed its genuine fears in the crash. Panic action and herding effect are all genuine fears to protect investors’ capital. COVID-19 created thrombosis in market liquidity as it did to its victims’ blood vessels that paralyzed the market cycle. In Folayan, Peterson & Kombe (2015), the paper analyzes the positive and negative variants of the same news about the Ebola virus, which originated in 2014 from Guinea in West Africa and jumped out to North America and Europe. This outbreak out of Africa is considered bad news. Filoviruses are known to science and medicine since 1967 (Marburg virus) and 1976 (Ebola virus). But Ebola did not get the attention it deserved thanks to the research sponsors and investors of pharmaceutical industry who did not consider it a priority, as the disease was not perceived as a threat or the investment in research of the Virus remunerative. “Global health experts have been saying that another pandemic whose speed and severity rivaled those of the 1918 influenza epidemic was a matter not of if but of when” (Gate, 2020).
Globalization and physical connectivity make any epidemic to spread massively (Achrekar et al., 2011; Brockmann and Helbing, 2013). Connectivity causes an outbreak (Jelinek et al., 2002). Thus, the implications for disease control and prevention are far-reaching. The “hemispheric or global spread of a series of diseases like Severe Acute Respiratory Syndrome (SARS), West Nile virus, and chikungunya” were cited in Nash et al., (2001). The work on treatments and vaccines for Covid-19 must gain further momentum. Research funding and attention of Pharmaceutical companies of various countries depend on how these countries were affected.

**IMPACT OF COVID-19 ON FINANCIAL MARKETS**

In Ozili & Arun (2020), the authors discuss how health crisis can translate to an economic crisis. According to the authors, the outbreak of the Virus led to social distancing, causing the closure of markets, activities, and businesses. Eventually, it affected supply-demand in all sectors, including financial markets except essential goods. International transactions came to a standstill. Periodical extension of lockdown days and restrictions on activities encouraged consumption of essential goods for survival and increased unemployment, and loss of jobs in some sectors. According to a sentiment analysis conducted using Twitter data for initial 4 days of lockdown in India, it was reported in Barkur & Vibha (2020) that Indians in general agree that lockdown paves the way for the exit route of the Virus and were angry that lockdown measure was a bit late. But I.T. and other activities performed digitally have experienced increased production of jobs. But the number of confirmed cases according to the authors did not have any significant effect on economic activities. In Conlon & McGee (2020), the paper explores whether Bitcoin acts as a 'safe haven' during the current COVID crisis and concludes that Bitcoin does not "provide shelter from turbulence in traditional markets."

In Baldwin & di Mauro (2020) and Mann (2020), the authors captured the impact of COVID-19 on the stock market and stated that smaller businesses will be more disrupted due to the pandemic. Stock markets capture investors' future expectations from an event based on fundamental value-drivers and velocity and volume of demand-and-supply. The relative changes in stock price index reveal which sectors and firms have potential to perform better in the future than others. The market reactions to the COVID-19 may be apparent in the way investors value the international trade and financial policies for firm value. Initially, investors’ negative sentiment affected multinational U.S. firms, having China exposure. As the Virus spread to Europe and the U.S., markets were agitated.

In McKibbin & Fernando (2020), using the modeling technique developed by Lee and McKibbin (2004) and extended by McKibbin and Sidorenko (2006), the paper explores seven scenarios in the context of studying the impact of COVID-19 on global economy and markets. The paper highlights that the state of health care systems in various economies have global ramifications. State of health care systems of individual nations is a global problem. The paper made preliminary estimates of the Cost of COVID-19 outbreak. Global cooperative investment in health care systems is highlighted. As the effects of pandemics are systemically critical to contemporary financial markets (Luo, 2020), health care of global population is a new obligation thrust on nations thriving on globalization. Schoenfeld (1982) researched at the firm level and opines that managers underestimate firm exposure to pandemic while publishing SEC-mandated risk factor filings though they have experienced a decrease in firm values during past as well as current virus episodes. The pandemic increased unemployment levels all over the globe and affected the values of bonds, commodities, and currencies across nations.

Investor behavior is captured in three phases in a Ramelli and Wagner (2020). In the first incubation phase in early January, some well-informed investors expected the transport sector to deteriorate during this early phase. Transportation and industry index and stock market index are positively connected. The World Health Organization (WHO) brought the onset of the Corona virus to the notice of investors after January this year. Discussions about the potential disruptive impact of the virus on
stock market started surfacing due to the increase in search trends of the virus leading to the plummeting of stocks of multi-national companies. As the performance of the tele-communication sector continued to be healthy, investors anticipated that the demand for services supporting work from homes like I.T. sector, Google, Face book, media firms, and FMCG, Food chain industry would boom. In the third phase starting from February 24, markets started to oscillate wildly. Since then, we witness panic selling in the stock market and panic buying in supermarkets. The market anticipation marks this phase for medium and long-term stock price movements. Firms with high leverage and cash deficit suffered due to overreaction (Ramelli and Wagner, 2020). COVID-19 is an eye-opener for investors and policymakers and the public that pandemic and natural disasters can considerably damage the global economy and financial markets. The impact of the pandemic on financial systems, cost of capital, insurance, pension planning, and policymaker's role would assume center stage in future research (Goodell, 2020).

In one study Albulescu (2020), the authors tested the impact of official news about COVID-19 on market volatility and reported that new cases reported outside china and news about increasing death cases have a positive impact on VIX. The outbreak of the Virus would continue to increase financial volatility. The persistence of the Virus has the potential to generate “a new episode of international financial stress.” Similarly, in Baker et al., (2020), the authors explored uncertainty in economic performance due to COVID. The paper identifies three indicators – volatility in stock market, News coverage on uncertain economic conditions and survey-based opinions on uncertainty in business expectation. These indicators can provide in real-time potential causes for uncertain business and economic conditions. Newspaper-based measures of uncertainty reflect what the media conjectures. Media influences peoples’ perceptions. Business expectation surveys are valuable indicators of firms’ threat perceptions in the short term. The paper depicted the possible impact of COVID-19 on the U.S. real GDP growth rate. According to the paper due to the impact of COVID-19, VIX has increased by about “500% from January 15, 2020, to March 31, 2020”.

From Hassan, Hollander, van Lent & Tahoun (2020), we can understand the implications of firm level exposure to pandemics including COVID-19. Using text-based measures of costs, benefits and risk exposure of the U.S based listed firms the study predicts which firms might gain or lose from the epidemic disease and which are most affected by the associated uncertainty due to outbreak across continents causing collapse of demand and disruptions in supply chain. Investors anticipate that firms which have successfully tackled earlier period pandemics would not be losers during COVID-19. The study is developed on text-based analysis using number of times the Virus is mentioned in quarterly earnings conference call the listed firms host with financial analysts.

The economic effects of COVID-19 on the global economy can be understood from various estimates of IMF, WHO, OECD and U.N. Department of Economic and Social Affairs (DESA). The pandemic crisis is challenging governments to implement monetary and fiscal policies that support credit markets to sustain economic activity amidst developing vaccines and protecting people. The combined effect of lockdown, oil-price war between Saudi Arabia and Russia, travel restrictions have led stock markets to their low esteem since 1987 and caused estimated loss of USD 113 billion in 2020 to airline industry. Airports in Europe report a loss of USD4.3 billion. Global recession for long periods anticipated with lower GDP growth estimates varying from country to country, which can fall lower than 10%. US GDP fell by 4.8% as per preliminary reports of Congressional research centre (CRS). It costs every month 2%-2.5% of global GDP as long as the crisis continues, and the economic costs of recession would be distributed unequally. Stock prices could tank 30% -40% in the ensuing future, giving investment opportunities to cash-rich nations and investors (Fernandes, 2020). Euro zone economy contracted by 3.8% with over 30 million people from European countries applied for state support.

The severely affected industries during the pandemic outbreak are units with massive investment in plant and machinery like Aviation and shipping as well as restaurants, hospitality, and tourism sector.
The least affected were food chains and FMCG sector manufacturing soaps and sanitizers. Labour intensive units obligated to pay employees without deploying in work are also affected. I.T. sector equipped with home-based jobs is protected as these firms continue to draw work from employees. The available information about U.S. Building and home products industries suggests that mortgage defaults triggered the global financial crisis in 2008 from the sub-prime crisis. COVID-19 crisis will not be different in these markets which would disturb even properly serviced loans also. Though most rated firms have adequate liquidity to crossover the crisis, energy and real estate sectors face prolonged problems during COVID-19 and post COVID-19. Due to migration of labour towards their home, labour crisis will be acute in markets like India. Banks have been facing problems from non-performing assets and the crisis further deepens due to erosion of capital and low deposits from customers. But state-owned banks in India enjoy protection from Government. According DESA, the COVID pandemic would shrink the global economy by up to 1% in 2020 instead of a previous growth projection rate of 2.5% as the pandemic is affecting supply chains and international trade during the global lockdown. During the crisis in 2009, the global economy had contracted by 1.7%, according to DESA. Governments of various countries need to take precautionary steps to restore the economy.

A decline in consumer-level expenditure in developed nations has reduced imports of consumer goods from developing countries hampering their economic prospects.

DESA opined that developing economies, particularly “those dependent on tourism and commodity exports, experience the effects of higher economic risks”. Global manufacturing activity could contract considerably and in turn, the tourism industry would suffer. Debt distress would occur due to capital outflows exceeding inflows. Governments should prefer to curtail public expenditure and ramp up spending on health care to fight the Corona virus. According to WTO, ‘world trade might fall by between 13% and 32% in 2020 as the COVID 19 pandemic disrupts’ regular economic activity and life around the world.

**BAD NEWS TRAVELS FASTER**

In Cinelli et al., (2020), the study addresses diffusion of information on COVID-19 using data sets from print media, audio, textual blogs and video forums and identifies information spreading from questionable sources. The study finds that information from both reliable and questionable sources get similar attention of the audience. Thus, markets respond to initial reaction of the investor from social media irrespective of its sources.

In Chen et al., (2014), Li et al., (2014) and Grossman and Stieglitz (1980), the authors found that the opinions expressed in articles predict stock returns and earnings surprises. It also found that language embedded in news and other media sources contains vital information for investors which is incorporated into stock prices (Davis, Piger, and Sedor, 2006). Thus, negative words predict low firm earnings and result in underreaction for a short span. When the negative words relate to the fundamentals, the reaction is largest and suggests that language embedded in news and other media sources would capture even hidden necessary information, and investors respond quickly by incorporating it into stock prices.

Tetlock (2010) believes that news has the potential to explain market price variances from time to time leading to market action as stock prices adjust readily to commonly known information. Heston and Sinha (2017), Uhl (2014) and Sprenger et al., (2014), demonstrated that information in the news could be deployed to predict both earnings yield and returns. For predictability to sustain for 3 months, the news is to be aggregated over a week and not daily. Thus proper aggregation is highlighted in the study. One reason for the effect of the news on prices is due to temporary sentiment or liquidity. By controlling for news, it shows that good news impacts stock prices within one week, but bad news can lower stock prices up to one quarter. Market absorbs negative news slowly as such it remains a
component of unincorporated news guiding future market direction. Hong and Stein (1999) show that the effect of analyst coverage on price is visible for past losers compared to past winners, and this is consistent with their hypothesis that firm-specific negative information penetrates slowly across retail traders during Bull Run. However, the news about the pandemic penetrated faster and deeper into the market even though the market was in a bull run due to herd effect.

The messages in wide circulation in Twitter or other micro blogging sites like Face book are considered fashionable and followed widely. Re-tweets are useful for analyzing content-based analysis. In Naveed et al., (2011), the paper states that bad news travels faster in the market. But in Coelho (2015), the authors show that the market cannot absorb bad news like bankruptcy events instantly. This asset pricing anomaly is explained in Hong and Stein (1999). The news about the pandemic did travel slowly from December 2019 to March 2020. In Kaminsky (1999), the World Bank, the paper examines the factors that trigger the financial markets and concludes that Investors overreact to negative news driven by herd instincts.

**IMPACT OF COVID-19 DUE TO THE PHENOMENON OF GLOBAL MARKET NETWORKING ON STOCK PRICES**

The dynamic linkage reactions of exchanges among four Asian countries – Taiwan, Singapore, South Korea, and Hong Kong were discussed in Masih, M.M. Abul and Masih, Rumi (1997) and concluded that the four markets are not mutually exclusive of each other and significant short-run linkages appear to impact each other. Similarly, Lau, Diltz & Apilado (1994) have observed exchange of information among Tokyo and New York stock markets (Agarwal, 2000) examined the ‘financial integration of capital markets’ in developing economies. Similarly, among dually listed stocks in more than one exchange, information transmission impacts the fortunes of one another, according to Bae, Cha and Cheung (1999). The transmission of COVID-19 information among various markets is due to panic market action and herding effect (Zhang et al., 2020).

Forss and Sarlin (2017) developed an algorithm for measuring sentiment-driven network risk to understand connectivity among firms and consequent risk dispersion. The authors have shown different results of which, when the risk model shows the highest quarterly risk value, the probable decline in the firm stock price in 70 days after a quarterly risk measurement is prominent. Though COVID-19 news started trickling into the market around December 2019, the impact occurred from March 11, 2020, when WHO declared it to be a global crisis. Balasubramanian et al., (2019) concur with the above findings. Thus, the clustering of “co-moving” stocks and even breaking of existing clusters cause “regime changes” in the market. The authors also present basic ideas on incorporating the effects of news and announcements on the stock price dynamics.

COVID-19 event is a typical herding effect. Herding effect in the financial market is identified as a tendency of investors to imitate others’ actions. The global financial markets are experiencing a herding effect after the pandemic news breakout. Markets experienced Panic liquidation due to the pandemic outbreak. At this stage, prices start tumbling as supply overflows the demand for stock due to fear of impending market collapse. People with Conservatism bias may have entered only after the news of COVID-19 has fully spread by April 2020, and hence they end up with snake bite bias! During this phase, market discounts, all kinds of bad news, and the bear market usually comes to a close before all bad news is made public furthering the bear effect. For long-term investors and seasoned investment managers relying on a stock’s fundamental strength, COVID-19 provided a great bargain due to the sudden tanking of value stocks. Experienced traders take the contrarian view characterized by speculation. Thus, the retail investor tends to lose many good chances for current investment. The herding bias is found to have an impact on risk-return tradeoff and influence stock price movement (Tan, Chiang, Mason & Nelling, 2008). The bias is characterized by overconfidence and fluctuations in
trading volume (Waweru et al., 2008) and differs from one sector to the other. Herding mentality is not limited to individual investors and may bug even institutional fund managers in panic situations.

We examined the implications of networking among exchanges in the case of the US, the UK, China and India using sentiment analysis tools as narrated in the methodology section.

**COSTS ASSOCIATED WITH RESUMING THE ECONOMY**

To regain normalcy and jumpstart the economy, firms incur substantial costs. Employers need the training to test and secure every employee in the workplace from the Virus. Employers need equipment and systems that test and ensure all employees joining back to service are not carriers of the Virus and immune. These costs are in addition to the costs involved in refurbishing plant and machinery. Employers need to re-evaluate their markets and competitive advantages. Employers need assurance that their employees protect them from Virus-related litigation from business rivals and from the general public. Budgets on Corporate social responsibility needs a further boost. These costs do not add up to making more profits but essential for remaining fit to work. Employers need to re-design systems and jobs to identify home-based works and others. Thus, in future investor sentiment is geared to measure costs of the pandemic impacting firm business model. Firms that can bounce back with minimum investment have a competitive advantage and add to the firm’s value in future asset pricing models. Gearing ratio in the capital structure accounts for its survival and guards against default risk unless financial institutions grant moratorium. But when firms raise debt funds from the public, the risk is acute and unguarded that need government intervention.

**DATA COLLECTION AND RESEARCH METHODOLOGY**

**DATA COLLECTION**

The paper aims to analyze the differential impact of the Corona virus on stock markets pre-lockdown and post-lockdown phase. A potential source for analyzing this impact is market news. However, market news is found to not immediately absorb and reflect rapid changes in the stock market in a short time span and takes a much longer time for absorption and reflection of the sentiment. Since the time period chosen for the study is short term, the data source for mining opinions is considered from Twitter (Teti et al., 2019) to perform and measure the variation in sentiments expressed about the Coronavirus during before the lockdown (January to March 31st) and during the lockdown (April 1st to May 16th) period.

Initially, for the pre-lockdown phase, 32000 tweets were extracted from January 1st to March 31st, 2020. The tweets were extracted based on the trending hashtag keywords that sparked debates on Twitter, i.e., #Coronavirus using the Twitter web API. Similarly, for the lockdown phase, 50000 tweets were extracted from April 1st to May 16th, 2020, based on the same keyword.

In order to analyze how the virus disrupted the stock markets, the search keyword was further narrowed down by augmenting the existing keyword with “#Coronavirus stocks” for both the above considered time periods (pre and during lockdown).

During the pre-lockdown phase, the number of tweets relevant to stock market (expressing investor opinions) was observed to be 10000 (high search volume indicating the trending nature of the disruptive impact on stock market). However, during the lockdown phase, the search volume reduced to 1000 reflecting that the lockdown played a crucial role in improving the stock market scenario.

However, to validate these findings and to examine the impact of the COVID pandemic event on the interconnectedness of global stock markets, we studied the stock markets of the USA (S&P 500),
the UK (London Stock Exchange), China (Shanghai Stock Exchange), and India (National Stock Exchange). For this, we used the daily closing prices of the above exchanges for both the pre-lockdown and lockdown periods collected from investing.com. We discuss the detailed methodology adopted for the sentiment analysis and for analyzing the stock market impact as under:

**RESEARCH METHODOLOGY**

**TWITTER SENTIMENT ANALYSIS OF CORONAVIRUS PRE AND DURING LOCKDOWN PERIOD**

We extract tweets and comments related to Coronavirus, and the discuss the methodology for pre-processing of the corpus, generation of word clouds, sentiment analysis and emotion analysis during pre and lockdown periods as under:

*Extraction of tweets and comments related to the Coronavirus*

The Twitter web API in R is used to extract relevant tweets about the virus. For assuring data security, an authentication token using ‘create_token()’ functionality is invoked providing access credentials only for downloading the tweets. The relevant English language tweets are then extracted (32000 during the pre-lockdown period and the 50000 tweets during the lockdown) before next steps of pre-processing.

*Pre-Processing the corpus*

The corpus is initially preprocessed before the analysis phase by removing special symbols, emoticons and irrelevant symbols with no insights. Stop words which do not directly have any implication for the term ‘COVID-19’ are then removed due to redundancy. Further, other steps like removal of duplicate retweets, stemming, lemmatization and tokenization is performed using the R tool text mining package ‘tm’.

After pre-processing this corpus of all COVID-19 related tweets, a wordcloud analysis to analyze the most frequent terms related to COVID-19 is performed on this overall corpus for both pre-lockdown and lockdown phases.

*Generating Word-clouds and Performing N-gram analysis*

The clean corpus is now converted into a Term Document Matrix (TDM). The TDM provides the terms and their cumulative frequency across all documents is computed for the word cloud analysis.

The predefined library 'word cloud' in R is used to generate the word cloud. A low frequency of 10 is found to generate a wordcloud with a large number of terms irrelevant to COVID-19. A high frequency of 100 largely limits the number of relevant terms which does not provide an insightful word cloud representation. Hence, to generate a more optimized word cloud with useful insights, a threshold frequency of term frequency=50 is chosen (Mishra, Raj & Pani, 2020). This visual representation provides a view of all the triggering keywords, which led to the opinions expressed on Twitter.

The wordcloud analysis provides the most frequently occurring words pertaining to COVID-19 but does not indicate the most probable sequence of words. In order to construct a probabilistic word distribution model that predicts based on a word, the occurrence of the next frequent word or next 2 frequent words, the N-gram analysis is performed. For the N-gram analysis, the filtered corpus of tweets related to impact of Corona virus on the stock market is adopted.
The most probabilistically occurring keywords are determined by unigram. Based on current word, the next frequent word is predicted by bigrams and similarly, the next two frequent words by a trigram analysis. The top 10 unigrams, bigrams and trigrams are thus obtained by the pre-defined library ‘RWeka’. The library contains the functionality ‘Ngram Tokenizer’ which specifies a control parameter to extract the required ngrams from the corpus by the generic R syntax:

\[
\text{NgramTokenizer}(\text{Corpus, Weka\_Control}(\text{min}=\text{number of grams, max}=\text{number of grams})\]

For instance, to extract bigrams (n=2) from corpus, the min and max parameters in above syntax are set to 2.

Similarly, unigrams and trigrams are extracted by specifying control parameter of 1 and 3, respectively. The results are enclosed below:

**Sentiment and Emotion Analysis**

In order to analyze the public perception about COVID-19 on the Twitter platform, sentiment analysis is performed from the tweets (both generic and specific to stock market). The sentiment polarity of the corpus (classifying text into different sentiment grades like “positive”, “negative” and “neutral”) is computed using a pre-defined “Sentiment Analysis” package in R tool. Further, to analyze the rationale behind the assigned sentiment categories, emotion analysis is performed in R tool using pre-defined “syuzhet” package. This classifies the emotions (Mohammad & Turney, 2013) using a dictionary-based approach.

While sentiment analysis and emotion analysis represent a more static picture of the public opinion, there is a need to visualize the sentiment distribution over time to capture changing opinions. Sentiment distribution charts are plotted for the chosen time period for both pre-lockdown and lockdown phases to capture variation in public opinion for both the periods.

**ANALYZING THE IMPACT OF CORONAVIRUS ON STOCK MARKET MOVEMENT**

The Coronavirus (COVID-19) impact on the global stock markets pre-lockdown and during the lockdown was analyzed to gauge the differential impact on stock markets. The effects on the stock market are reflected in the extent of connectedness between the different country stock markets before imposing the lockdown and during the lockdown period.

For measuring the connectedness, the mutual spillover effects of the major stock markets of countries USA, UK, China, and India were estimated using the Vector Auto-regressive model (VAR) model for both the different periods.

Based on the spillover values, the connectedness graph or adjacency graph (minimum spanning tree) was plotted between the different country stock markets pre and during the lockdown (Phan & Narayan, 2020) using the ‘frequencyConnectedness’ package in R. The results of the spillover matrix and the minimum spanning tree graphs are thus illustrated below.

The methodology is illustrated below in Figure 1:
RESULTS

This section presents the sentiment analysis and stock market analysis results.

GENERATING WORD CLOUD FOR TWITTER AND PERFORMING NGRAHM ANALYSIS

The word cloud in Figure 2 illustrates that the words “covid”, “january”, “virus”, “deaths” and “pandemic” are most frequently used by people since the virus was first detected in January with potential disruptive impact on society and businesses.

Further, the word cloud is illustrated during the lockdown period below in Figure 3:
The lockdown period illustrates an improvement in public safety with emphasis laid on testing, self-quarantining and using precautionary measures like face masks and shields. This is evident from the frequently occurring words in wordcloud: ‘safe’, ‘test’, ‘self-quarantine’ and ‘mask’.

Further, the top 10 unigrams, bigrams and trigrams during the above period (for both pre-lockdown and lockdown) were:

![Top 10 Unigrams](Image)

In Figure 4, the top 10 unigrams are: ‘covid’, ‘cases’, ‘pandemic’, ‘quarantine’, ‘government’, ‘schools’, ‘lockdown’, ‘businesses’, ‘economy’ and ‘mask’. This indicates that the COVID-19 has a disruptive impact on government, schools, economy and businesses and there is a need to enforce a lockdown and wear masks as safety precaution.

Further, a bigram analysis of the same corpus is performed illustrated below in Figure 5:
The top 10 bigrams are observed to be: “covid cases”, “about covid”, “dangerous virus”, “the pandemic”, “tested positive”, “dipping shares”, “covid patients”, “covid patients”, “crashing economy” and “covid conspiracy”. The analysis confirms the above unigram findings and further qualifies the extent of disruptive impact on the stock market by the terms “dipping shares” and “crashing economy”. Thus, COVID-19 is reported to have led to crashing economy and plummeting stock market shares which needs to be investigated in detail by performing stock market analysis.

Further, to validate the above unigram and bigram analysis, a trigram analysis is also performed on the corpus to identify the most frequently occurring 3-gram words capturing the impact of COVID-19.

The top 10 trigrams in ascending order of frequency (in Figure 6) are observed to be: “cases and deaths”, “for us vaccines”, “stock market bear”, “flexible working conditions”, “spread of covid”, “biden trump covid”, “the covid pandemic”, “positive for covid”, “new covid cases” and “due to covid” which corroborate that the COVID-19 has accentuated the bearish nature of stock markets and due to its widespread pandemic characteristics, all businesses and companies have enforced flexible working conditions like “work from home”.

The Twitter Sentiment and Emotion Analysis results are illustrated below:
TWITTER SENTIMENT AND EMOTION ANALYSIS RESULTS

SENTIMENT AND EMOTION ANALYSIS PRE-LOCKDOWN

![Sentiment Analysis Pre-Lockdown](image)

**Figure 7.** Sentiment Analysis of Opinions About Coronavirus Pre-Lockdown

Figure 7 illustrates a dismal picture of the opinions about the Coronavirus during the pre-lockdown phase. Negative sentiment tweets are found to be in high numbers (295) outnumbering the number of positive tweets (170). Neutral sentiment tweets are 267 in number. This reflects the disruptive impact of COVID-19 on the society and calls for extensive precautionary measures.

![Pre-lockdown Emotion Analysis](image)

**Figure 8.** Emotion Analysis Expressed About Impact of Coronavirus on Stock Market Pre-Lockdown

The emotion analysis chart in Figure 8 illustrates that negative emotions are pre-dominant due to the potential impact of the virus on stock market. However, the positive emotion ‘Anticipation’ is high indicating that people are expecting a solution to the disease (vaccine or antidote) to cure the virus.

In order to analyze the varying opinions overtime, i.e., during the pre-lockdown period, the sentiment polarity is plotted versus the corresponding date using the plot() function and is illustrated below in Figure 9:
The sentiment represents a completely gloomy picture of the Coronavirus. Initially, the sentiment score is highly negative during the 2nd week of February due to new cases in China and South Korea. This was accentuated by the spread of the pandemic to Italy and further to other European countries, the plummeting global stock market prices. Subsequently, the increasing occurrence of virus positive cases in India and the increasingly damaging impact of the virus demanded for a world-wide lockdown phase.

SENTIMENT AND EMOTION ANALYSIS DURING THE LOCKDOWN

Further, during the lockdown period, the results are illustrated in Figures 10-12:

The opinions about Coronavirus are analyzed to know the sentiment during the lockdown period. This is depicted in a sentiment polarity chart illustrated in Figure 10 with an increase in positive number of tweets to 370 outnumbering the negative tweets (295). The number of neutral sentiment tweets
were observed to be 450. This result confirms that the lockdown was needed to take the situation into control and the higher number of positive tweets indicates that the damage to businesses and society was partly mitigated as employees (most of them use Twitter) would be safe.

![Distribution of Emotions During Lockdown](image)

**Figure 11. Emotions Distribution During the Lockdown**

The emotion distribution pie chart in Figure 11 portrays the various emotions during the lockdown. The distribution is found to change with people expressing on Twitter more positive emotions like positivity (20%), joy (9%), surprise (4%), trust (14%) and anticipation (8%) (indicating the lockdown enforced helped to an extent to contain the virus. However, there are still skeptical reactions to the enforcement of lockdown with 45% of people (aggregate of all negative emotions) displaying anger (6%), fear (4%), sadness (8%), disgust (9%) and negativity (18%). This indicates that there is still some work to be done to reduce the disruptive impact of the pandemic.

The sentiment distribution of the opinions during the lockdown period illustrated in Figure 12.

![Sentiment Distribution of Corona Virus During Lockdown](image)

**Figure 12. Sentiment Distribution of Tweets About the Coronavirus Over the Lockdown Period**
Figure 12 depicts a net positive opinion with peaks on April 23rd, April 27th, May 6th, and May 14th. This is in stark contrast to the pre-lockdown sentiment distribution chart illustrated in Figure 7, where the sentiment was highly negative and varying between high to moderate negative.

Though there is a sharp increase in a spike of the positive cases, the sentiment distribution is governed more by the recovery rate and several deaths. Since the recovery rate has improved, and deaths have reduced during the lockdown, the sentiment distribution is between moderately favourable to positive during this period.

This result affirms that the imposition of lockdown was necessary for controlling the number of deaths due to the Coronavirus and also for minimizing the disruptive impact on the stock market.

Further, to analyze the impact of the virus on inter-connectedness between stock markets, the following analysis is adopted as a validation to the above sentiment analysis:

**IMPACT OF THE CORONAVIRUS ON STOCK MARKET (PRE-LOCKDOWN AND DURING LOCKDOWN)**

The ‘frequencyConnectedness’ package in R computes the influence of one country stock market over the other by constructing a mutual influence table termed as “Spillover matrix” for all the countries considered in the analysis. The “Spillover matrix” in Table 1 is obtained from the pre-defined function ‘spilloverDY09’. This provides the variance error spillover matrix in Table 1 for the four stock markets (USA, UK, China, and India) during the pre-lockdown phase:

<table>
<thead>
<tr>
<th>Total Spillover</th>
<th>USA</th>
<th>UK</th>
<th>China</th>
<th>India</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>95.56</td>
<td>1.69</td>
<td>0.32</td>
<td>2.43</td>
</tr>
<tr>
<td>UK</td>
<td>32.01</td>
<td>45.84</td>
<td>5.00</td>
<td>17.15</td>
</tr>
<tr>
<td>China</td>
<td>20.10</td>
<td>5.21</td>
<td>72.15</td>
<td>2.54</td>
</tr>
<tr>
<td>India</td>
<td>11.96</td>
<td>21.56</td>
<td>3.04</td>
<td>63.45</td>
</tr>
</tbody>
</table>

In the matrix, the diagonal values indicate the component of variation in a single country stock market which does not “spillover” to other countries. This component is attributed to events in a country stock market which are only internal to that country and have no ramifications for other countries. For instance, in Table 1, row 3 column 3, the diagonal value 45.84 indicates that 45.84% of the variation in UK stock market is due to internal events which do not have a network effect on other countries.

The numeric spillover values provided in the other non-diagonal elements of the table are indicative of the influence of variation in a single country stock market to the other country stock markets.

For instance, in the first row, of a total of say 100 units (percentage), a variation in US stock market is found to influence the UK stock market to the extent of 1.69 units (1.69%), Chinese stock market by 0.32 units (0.32%) and Indian stock market by 2.43 units (2.43%). The remaining component which does not spillover to other country stock markets and only internally influences US stock market movement is = 100 units - (1.69+0.32+2.43) = 95.56 units (95.56% inherent to US).

The positive spillover values indicates a positive influence to other country stock markets. indicating that any variations in the USA stock market influences the UK, China, and Indian stock markets positively.

Similarly, UK is found to be the next influential stock market with a spillover of 32 units to the USA, 5 units to China, and 17.15 to India, indicating that variance in the UK stock market is penetrating to the USA, India, and China. China has a high spillover to the USA, while India is found to spillover to the UK.
Further, from the net spillover computed above to determine the connectedness of stock markets, two minimum spanning trees are plotted before and during lockdown phase. The purpose of the minimum spanning trees is to graphically represent the mutual influence of stock markets.

In the minimum spanning tree, each country is represented by a small node with country name. If the topology of the nodes is arranged in such a way that there are no central nodes and all nodes are connected by a line, this implies that no stock market of any country has undue influence (or spillover) to another country stock market.

However, if any country node is found at the center with lines connected to it from other nodes, this implies the central node is the country influencing other countries.

Following this notation, the minimum spanning tree (Zhang et al., 2020) is plotted for the stock markets illustrated below in Figure 13:

![Minimum Spanning Tree (MST) of Stock Markets Pre-Lockdown](image)

From the minimum spanning tree above, it is observed that the node pertaining to USA stock market is at the center of all nodes and is connected from all the nodes. This topology indicates that during the pre-lockdown phase, USA stock market was at the center of action followed by UK and events in these stock markets are found to influence the stock market movement in Asian countries like India and China indicating that no similar event has occurred which equally influenced all the stock markets.

During the lockdown period, the net spillover matrix is illustrated in Table 2 as follows:

<table>
<thead>
<tr>
<th>Total Spillover</th>
<th>USA</th>
<th>UK</th>
<th>China</th>
<th>India</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>97.94</td>
<td>0.95</td>
<td>1.10</td>
<td>0.01</td>
</tr>
<tr>
<td>UK</td>
<td>22.65</td>
<td>72.06</td>
<td>5.28</td>
<td>0.01</td>
</tr>
<tr>
<td>China</td>
<td>21.18</td>
<td>1.32</td>
<td>77.47</td>
<td>0.03</td>
</tr>
<tr>
<td>India</td>
<td>35.57</td>
<td>11.75</td>
<td>12.30</td>
<td>40.37</td>
</tr>
</tbody>
</table>

The spillover of USA to UK, China, and India is found to considerably decrease in Table 2 with only 0.95 units to the UK, 1.1 units to China, and 0.01 units to India. Similarly, the UK has also reduced its spillover to China and India considerably, as seen above.
The minimum spanning tree is plotted for the stock markets during the lockdown period illustrated below in Figure 14:

![Minimum Spanning Tree (MST) of Stock Markets During the Lockdown](image)

From the minimum spanning tree in Figure 14, it is found that there are no central nodes in the topology and all nodes are at the corners connected to each other. This corroborates the above result illustrating that no single country is in the center of the action and there is no dominating central node influencing other nodes. This decrease in the domination of a single country stock market indicates that the Coronavirus lockdown equally influenced all the country stock markets that have had a devastating impact on businesses across the globe. The rate of recovery of the respective countries concerning the Coronavirus determines the stock market recovery independently for each country. Thus, the Coronavirus is found to be an equalizer for all the developing and developed economies, and the imposition of lockdown is also found beneficial in terms of reducing stock market spillovers which could have further disrupted the businesses, particularly, of multi-national companies listed commonly on all the four stock exchanges.

**IMPLICATIONS AND RECOMMENDATIONS**

The main contribution of the paper is to capture the change in opinions among investors before and after enforcing a lockdown and to validate the differential impact on the stock market interconnectivity for both phases. Twitter sentiments appear to show that market leadership is shattered due to the pandemic. The results of the stock market network analysis reveal that prior to lockdown, the US markets were at the centre of action of the stock market and led the global market directions as shown in the minimum spanning tree in Figure 13. This single country domination however was not visible with the onset of the COVID-19 era and lockdown phase as depicted in the minimum spanning tree in Figure 14. This implies that until the US market stabilizes, other markets networked in the globalization seem to follow country specific conditions to gather momentum in their respective markets. Differences in policy configurations among countries result in reduced co-ordination among nations to tackle the aftereffects of the pandemic. This would pave the way for strained relation among developed and developing economies and among northern and southern Euro zone. This might cause strained relation in alliances leading to reestablishing global leadership.
Perhaps to retain global leadership, the US seems to have opted for protecting their economy first in the run for developing a vaccine. Delayed action for lockdown evidences this tendency.

According to CRS report global economic recession led foreign investors in Asian countries to pull out over USD 26 billion (over 16 billion from India itself) causing collapse of these markets. Reduced China exports in agricultural products, automobile parts, computers, cell phones, toys and medical equipment impacted supply chain in Japan and other countries. Despite reduced foreign funds flow, China, India and Indonesia may record mild positive economic growth in 2020 among the global nations. Certainly, Coronavirus marks a new era in redefining global economic leadership alliances.

However, due to the uncertainty in the stock market, investors are recommended the following in an Indian stock market context (Jain, 2020):

- Therefore, investors must be cautious and invest based on the analysis of individual stock performance. It is recommended to keep a watchful eye for large capitalization and selected mid-capitalization firms and proceed in a cautious and staggered manner.
- It is observed that COVID-19 than other sectors relatively less disrupts some sectors. Sectors of chemicals, agricultural sector, and Fast-Moving Consumer Goods (FMCG) stocks have a higher recovery rate.
- It is advised to invest in Blue-chip private sector banks. The insurance sector is also eyeing an upward trend due to people's awareness about healthcare and life eventualities in the post-Covid scenario.
- Thus, it is preferred that investors focus on investing and trading on “quality” shares on a short-term basis. The investment must have a minimum time horizon of 1 year to 2 years given the current scenario's inherent risk.

CONCLUSION AND FUTURE SCOPE FOR RESEARCH

The paper has attempted to measure the Corona Virus pandemic's impact on the economy and stock market. The analysis is divided into two phases:

The first phase involves adopting a social media Twitter-based sentiment analysis to initially capture the disruptive impact of COVID-19 pre-lockdown. The highly negative sentiment polarity in Twitter reflected the damaging impact on the stock market. The results thus enforced the need for imposing a lockdown. The consequential impact of lockdown on COVID-19 and in turn, on the stock market was then analyzed on Twitter, which showed a marginal improvement though still indicative of a dangerous phase for investors. The second phase of the study corroborated the insights.

The second phase involves validating the Twitter sentiment analysis results by analyzing the Coronavirus (COVID-19) impact directly on global stock markets pre-lockdown and during the lockdown to gauge the differential impact on stock markets. The impact was measured in terms of the extent of connectedness between the different country stock markets before imposing the lockdown and during the lockdown period for which the mutual spillover effects of the major stock markets of countries USA, UK, China, and India were estimated using the Vector Auto-regressive model (VAR) model for both the different periods. While pre-lockdown analysis showed the supremacy of the US stock market over other countries, the lockdown period demonstrated a de-centralized impact with a decrease in the US stock market's domination. Further, the lockdown period also helped reduce stock market spill overs, indicative of a minor improvement in the global stock market scenario. Thus, the COVID-19 is affirmed to be an equalizer for all the developing and developed economies. Further, in the current scenario, recommendations are provided to the investors to be more cautious and wary of the fluctuating stock market scenario.

The study's time period could be extended to examining the ramifications of COVID-19 on the stock market in the post-lockdown period and on a longer horizon of 1-2 years. Further, the study was limited
to the top developed and developing stock markets and could be extended to analyzing the spill overs in different country stock markets. The impact of COVID-19 on other secondary markets like derivatives and bond markets can be investigated. For sentiment analysis, techniques like maximum entropy-based Joint-Aspect-based Sentiment Analysis Technique (JABST) (Tang et al., 2019) can be leveraged for a more probabilistic prediction of the varying sentiment.

It is thus recommended that the comity of nations should come together to contain the virus from causing further disruptive impact on the society.
REFERENCES


