# Major Sudden Risk Shocks and US Stock Market Volatility Evidence from COVID-19

Kepeng Liu[1, *]

[1] *Reading Academy, Nanjing University of Information Science and Technology, Nanjing, 210044, China*
[*] *Corresponding author. guanghua.ren@gecacademy.cn*

**ABSTRACT**
The occurrence of major emergencies will impact the stock market, and the short-term fluctuation and rebound of stock prices can be estimated through relevant data to predict the impact of emergencies on the stock market. By employing a GARCH (1,1) model, the empirical analysis of the impact of Corona Virus Disease 2019 (COVID-19) on the US stock market was conducted by examining the maximum price, closing price in the stock market, and the number of newly confirmed cases. It was shown that the COVID-19 epidemic did not cause fluctuations in the US stock market in the long term. However, from the estimation results of the highest price, the new diagnosis reduces the volatility of the return series of the highest price of the stock market.

***Keywords:*** *US stock market, COVID-19, ARMA, ARMA-GARCH.*

## 1. INTRODUCTION

It is well known that the US stock market is affected by many factors and the impact of major unexpected risks on the US stock market is particularly obvious. Many people who lack expertise in the financial field are often at a loss when they encounter such events. Then it is particularly important to find the impact and relationship of major unexpected risks on stock market fluctuations because this enables investors to avoid risks and reduce losses in time when they encounter such shocks. COVID-19 refers to pneumonia caused by novel coronavirus 2019 infection. Since December 2019, several cases of pneumonia of unknown cause with a history of exposure to seafood markets in South China have been found in some hospitals in Wuhan, Hubei Province, China. And then COVID-19 cases were subsequently diagnosed around the world.

**Table 1.** COVID-19 data statistics as of date October 2nd, 2021

| No. | Country | COVID-19 cases count | |
|---|---|---|---|
| | | Confirmed | Death |
| 1 | United States | 44,443,405 | 718,984 |
| 2 | India | 33,789,398 | 448,605 |
| 3 | Brazil | 21,445,651 | 597,292 |
| 4 | United Kingdom | 7,841,625 | 136,789 |
| 5 | Russia | 7,535,548 | 208,142 |
| 6 | Turkey | 7,182,943 | 64,264 |
| 7 | France | 7,018,367 | 116,759 |
| 8 | Iran | 5,601,565 | 120,663 |
| 9 | Argentina | 5,258,466 | 115,225 |
| 10 | Spain | 4,961,128 | 86,463 |

Notes: The ten countries most affected at the moment. To get the newest data on the number of confirmed cases in each country, please visit https://coronavirus.jhu.edu/map.html，the website is developed by the Center for Systems Science and Engineering at Johns Hopkins University.

Because of the widespread of the virus and its low mortality, during asymptomatic periods, it is difficult to detect. By the time people noticed and started to prepare for prevention, there were already large numbers of patients, especially in the United States. According to the relevant epidemic data provided by World Health Organization, it shows a sharp rise in daily new diagnoses and deaths in the United States since the first case was reported in mid-January 2020. As of August 31st, 2021, nearly 40 million people in the United States had been diagnosed and about 650,000 Americans died of COVID-19.

By exporting and analyzing the k-charts of the S&P 500 in the United States from approximately three months before the first announcement of COVID-19

Copyright © 2022 The Authors. Published by Atlantis Press International B.V.
This is an open access article distributed under the CC BY-NC 4.0 license -http://creativecommons.org/licenses/by-nc/4.0/. 

patients to August 31st, 2021, we can see that the overall trend of the stock market is still a slow upward trend, but there were several considerable decreases in 2020, most notably from early February to the end of March in, the middle of August to mid-September and mid-October to early November. By referring to the US epidemic data, we know that early February was shortly after the United States reported its first COVID-19 case. In mid-August, the daily rate of new cases in the United States began to rise sharply, and since early October the rate has been increasing and the daily number of new cases has exploded.

Most existing literature on unexpected risk shocks on the stock market mainly focus on stock market volatility in countries like UK or China but a few on the US, especially in the period before and after the epidemic when the US stock market has been significantly affected. An emergency can cause fluctuations in a country's stock market. Recent research showed that on the day of the Brexit referendum, there was a negative impact on the stock markets of major countries in the world, among which the decline in the European Union was the largest, indicating that Brexit had the most profound impact on the European Union, while it had the least impact on the Chinese stock market[1]. At that time, since countries outside the European Union (EU) such as the United States and China were not greatly affected by the Brexit referendum, which was mainly aimed at the EU, then we could not intuitively study the impact of emergencies on the US stock market. And now COVID-19 has spread around the world, with the United States being most affected. The research found that in particular, sentiment analysis using big data from social media is an excellent source of information for investors to decide on investment strategies when the stock market is hit by unexpected shocks like COVID-19. As the COVID-19 pandemic continues, it is still impossible to predict the extent of its global impact [2]. We can believe that COVID-19 did have a big impact on the US stock market, but this literature has some limitations because it is based on the data from January 2020 to May 2020, In other words, we need to combine the most recent year's data to come up with a conclusion that better explains the trend and volatility of the US stock market.

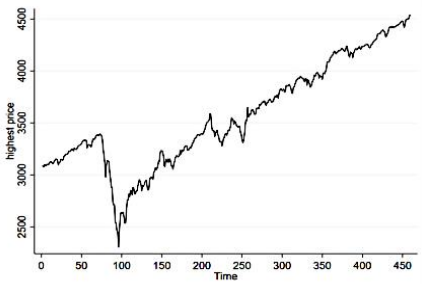
In our paper, since most relevant literature all points out that COVID-19 has had varying degrees of impact on the global economy, especially in US, and recent research may be out of date, that is, it may no longer be applicable to the current situation of the US stock market, we will take COVID-19 as an example of a major sudden risk shock o study its impact on U.S. stock market volatility. In order to exclude the impact of the trade war between China and the US happened shortly before the pandemic on the stock market, a period from November 4, 2019, to August 31, 2021, will be included in our study.

## 2. DATA AND METHODOLOGY

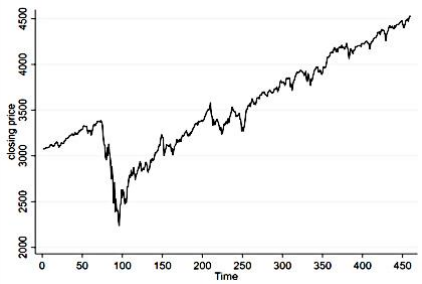
### *2.1. Data and Summary Statistics*

The study used the S&P 500 from November 4th, 2019 to August 31st, 2021 provided by standard & poor's (index data can be obtained at http://quote.eastmoney.com/gb/zsSPX.html). The S&P 500 is a stock index of 500 publicly traded companies in the United States. The stock index is created and maintained by Standard & Poor's. All companies covered by the S&P 500 are listed on major U.S. exchanges, such as the New York Stock Exchange and Nasdaq. The S&P 500 includes more companies than other indexes, such as the Dow Jones, so the risk is more diversified and reflects broader market changes.

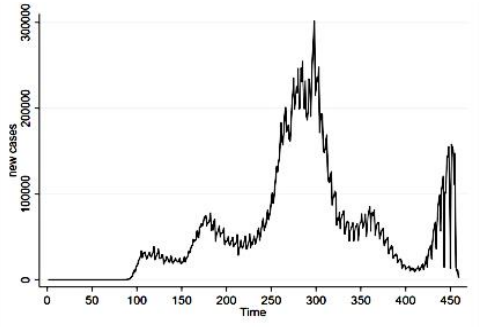
Keeping November 4th, 2019 as the first day in our study, we collected the daily data for the highest and closing prices of the S&P 500 during this period, and among them, we focused more on the daily highest price. In addition, we also collected the relative data on COVID-19 in the United States since November 4, 2019, and made them into line charts, which would be used later in studying its relation to the volatility of the U.S. stock market.




**(a)**




**(b)**




**(c)**

**Figure 1.** The four S&P 500 indices and the daily number of new cases in the United States during the period from November 4th, 2019 to August 31st, 2021. (a) The highest price, (b) Closing price, (c) New cases in the United States.

## *2.2. Basic Model and Stability Test*

Recent research showed that the Autoregressive Integrated Moving Average Model (ARIMA) for finance was used to explain the potential connection between two things. Box-Jenkins's method is a perfect and accurate algorithm for analyzing and predicting time series data, and its commonly used models include Autoregressive model (AR), Moving Average model (MA), Autoregression Moving Average Mixed model (ARMA), and Autoregressive Integrated Moving Average model (ARIMA) [3,4]. Based on these models, our final aim was obtaining Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model to study the exact connection between the COVID-19 pandemic and the US stock market. Another study had shown that Unit root test is the preferred method to estimate whether time series is stationary or not and Logarithmic transformation and difference help to alleviate the heteroscedasticity problem. Auto-correlation function (ACF) and partial autocorrelation function (PACF) graphs were used to identify the orders of the ARIMA model[5]. Since unstable data mostly are difficult to find out the correlation, so we first need to test whether the data are stable. Recent research mentioned that the stationarity of time series is the statistical characteristic of mean and variance of a series with time. If both of them are constant over time, then we said the series is stationary, otherwise, the series is a non-stationary process [6]. ACF is defined by $\{\rho\}$[7]. Here $\rho$ refers to the auto-correlation coefficient:

$$\rho_k = \frac{\mathrm{Cov}(x_t, x_{t-k})}{\sqrt{\mathrm{Var}(x_t)Var(x_{t-k})}} = \frac{\mathrm{Cov}(x_t, x_{t-k})}{\mathrm{Var}(x_t)} \quad (1)$$

And since the correlation coefficient is defined as:

$$\rho_{x,y} = \frac{\mathrm{Cov(X,Y)}}{\sqrt{\mathrm{Var(X)Var(Y)}}} = \frac{\mathrm{E}[(\mathrm{X}-\mu_x)][(\mathrm{Y}-\mu_y)]}{\sqrt{\mathrm{E}(\mathrm{X}-\mu_x)^2\mathrm{E}(\mathrm{X}-\mu_y)^2}} \quad (2)$$

Then the sample correlation coefficient is the consistent estimate of the correlation coefficient, that is:

$$\hat{\rho}_{x,y} = \frac{\sum_{t=1}^{\mathrm{T}}(x_t-\bar{x})(y_t-\bar{y})}{\sqrt{\sum_{t=1}^{\mathrm{T}}(x_t-\bar{x})^2 \sum_{t=1}^{\mathrm{T}}(y_t-\bar{y})^2}} \quad (3)$$

In (1) $Cov(x_t,x_{t-k})$ is the covariance of $x_t$ and $x_{t-k}$. $x_t$ refers to the t-th number in the sample. $Var(x_t)$ stands for the variance of $x_t$. In (2) $\mu_x$ is the expected value of the random variable $x$. In (3) $\bar{x}$ refers to the sample mean.

Then we would apply Unit Root Test to verify the stability of our data. Its model is defined as:

$$x_t = c_t + \beta x_{t-1} + \sum_{i=1}^{p-1} \varphi_i \Delta x_{t-i} + e_t \quad (4)$$

In (4), t = 1, 2, ... is a unit root process.

In Stata, we chose Augmented Dickey-Fuller (ADF) test to test the stability. The output of the ADF test includes the ADF statistics to test the coefficient of lag variables and the critical values required for the test (1%, 5%, 10%). If the coefficient is significantly non-zero and is less than zero, then the null hypothesis containing the unit root is rejected and the alternative hypothesis is stable will be accepted. Based on these we established the following hypothesis: $\mathrm{H}_1^{\mathrm{S}}$

$\mathrm{H}_1^{\mathrm{S}}$: The data we had collected were unstable.

We investigated the following null and alternative hypotheses to test the hypothesis $\mathrm{H}_1^{\mathrm{S}}$.

$$\mathrm{H}_0: \beta = 1 \text{ vs } \mathrm{H}_1: \beta < 1 \quad (5)$$

After verifying the stability of closing price and highest price in Stata, we summarized the main information as the following form:

**Table 2**. Results of stability test

| Indices | Parameters | | | | |
|---|---|---|---|---|---|
| | Test statistic | Critical value 1% | Critical value 5% | Critical value 10% | P-value |
| Logarithm series-closing price | -2.07 | -3.98 | -3.44 | -3.13 | 0.57 |
| Logarithm series-highest price | -2.25 | -3.98 | -3.42 | -3.13 | 0.46 |
| Logarithm return rate series-closing price | -14.68 | -3.98 | -3.42 | -3.13 | 0 |
| Logarithm return rate series-highest price | -13.27 | -3.98 | -3.42 | -3.13 | 0 |

Notes: The Logarithm and Logarithm return rate series. In Stata, we calculated test statistic critical value 1%, critical value 5%, critical value 10% and p-value respectively.

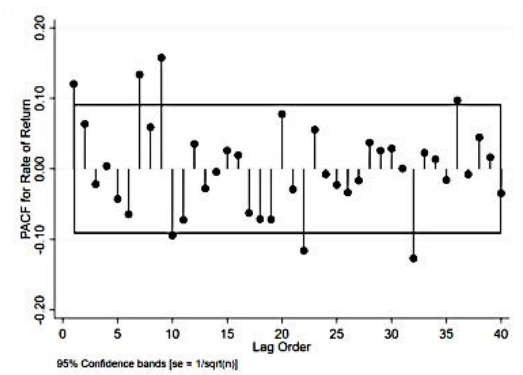
Note that in the form the p-value for logarithm series of the closing and highest price were quite great so that we would accept the null hypothesis and chose to use the logarithm return rate series, whose p-value were 0 and this meant that the logarithm rate series for closing and highest price are stable. And our next step was to figure out the order of ARMA in Stata so that we could construct ARMA.

## 2.3. Orders of AR and MA

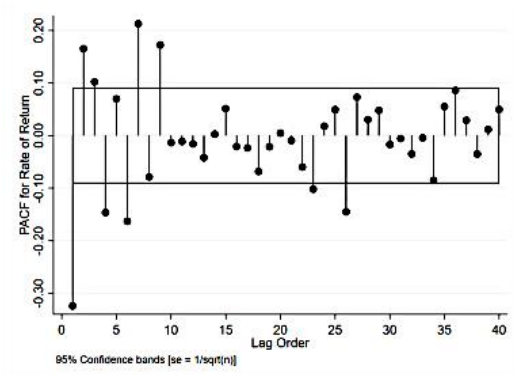
With the property that the auto-correlation coefficient of a sample is a consistent estimate of the auto-correlation coefficient, and follows a normal distribution with a mean of 0 and variance of 1/T, we could determine the orders. For stationary data，AR(p) is set as follows:

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \ldots + \varphi_p x_{t-p} + a_t \quad (6)$$

In (6), $p$ is a non-negative integer. $\varphi_0$ is a constant term. at is assumed to be a random error with the mean equal to 0 and the standard deviation equal to $\sigma$ which is assumed to be constant for any $t$. In order to reduce computation and for convenience, we would determine the order as small as possible. By applying PACF we got the following graphs:




**(a)**



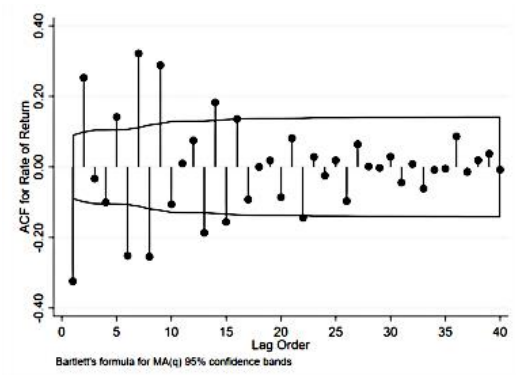

**(b)**

**Figure 2.** The PACF for rate of return. (a) Closing price, (b) highest price.

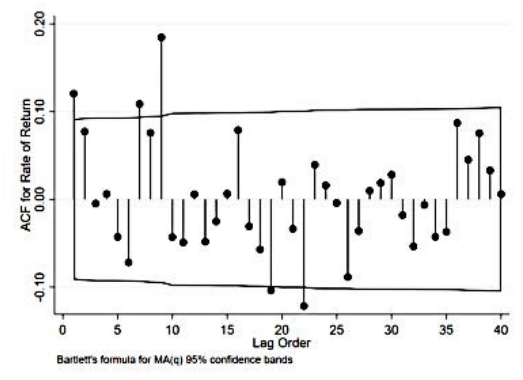
According to the graphs we could determine that for AR, the orders of these two indices were both 1. MA is set as follows:

$$x_t = c_0 + a_t - \theta_1 a_{t-1} + \ldots + \theta_q x_{t-p} + a_t \quad (7)$$

We would apply ACF to determine the order of MA.




**(a)**




**(b)**

**Figure 3.** The ACF for rate of return. (a) Closing price, (b) highest price.

These two graphs indicated that the orders of logarithm rate series for MA were also 1.

### *2.4. ARMA and ARMA-GARCH*

ARMA(p,q) contains AR and MA, and it is set as follows:

$$x_t = \varphi_0 + \sum_{i=1}^{p} \varphi_i x_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i} \quad (8)$$

In (8), $\{a_t\}$ is a white noise sequence, $p$ and $q$ are non-negative integers. AR and MA both are special cases of ARMA and we had got that the order of both the closing sequence and the highest sequence was 1, which meant that in ARMA, $p$ and $q$ were both equal to 1.

Based on ARMA, we needed to construct GARCH. In fact, GARCH comes from ARCH. Known as a new theory, the ARCH model has developed rapidly in recent years and has been widely used to verify the rule description in financial theory and the prediction and decision of the financial market. The basic idea of ARCH model is that under the previous information set, the occurrence of noise at a certain time is normal distribution. The mean of the normal distribution is zero and the variance is a quantity that changes over time (i.e., conditional heteroscedasticity). And the variance over time is a linear combination of the squared noise values of a finite number of past terms (i.e., auto-regression). Thus an Autoregressive conditional heteroscedasticity model is constructed and it is set as:

$$y_t = x_t'\beta + \varepsilon_t \quad (9)$$

In (9), yt is the dependent variable, xt is the independent variable and εt is the perturbation term. Denote the conditional variance of the perturbation term εt as $\sigma^2{}_t \equiv \mathrm{Var}(\varepsilon_t \mid \varepsilon_{t-1}, \ldots)$. The subscript t of $\sigma^2$ indicates that the conditional variance can change over time. Suppose $\sigma^2$ depends on the square of the perturbation in the previous period:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 \quad (10)$$

which is exactly ARCH(1). More generally, suppose A depends on the square of the first p perturbation terms:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \ldots + \alpha_p \sigma_{t-p}^2 \quad (11)$$

In ARCH(p), if p is very large, many parameters need to be estimated and the sample size will be lost. GARCH, recently proposed by Bollerslev(1986), reduced the parameters to be estimated and predicted the future conditional variance more accurately. On the basis of ARCH model, the Autoregressive part of $\sigma^2{}_t$ is added. Hence we chose to construct ARMA-GARCH. GARCH(p,q) is defined as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \ldots + \alpha_p \sigma_{t-q}^2 + \gamma_1 \sigma_{t-1}^2 + \ldots + \gamma_p \sigma_{t-p}^2 \quad (12)$$

Note that ARCH(GARCH) models can only be constructed if there is conditional heteroscedasticity. Naturally, we need to test whether the conditional heteroscedasticity (i.e., ARCH effect) exists. The most accessible method is to estimate the GARCH(or ARCH) to see whether the coefficients of α or γ in the variance equation are significant so that the significance tests will be included in subsequent results and discussion. Recent research showed that despite the wide variety of GARCH models, GARCH(1,1) is generally regarded as the most popular model in GARCH-type models in many empirical applications. It is convenient to calculate and widely used by experts to model the volatility of daily returns.[8]. Hence in our study, we would construct GARCH(1,1).

## 3. RESULTS AND DISCUSSION

### *3.1. Results*

In Stata, we took ARMA as the mean equation and further introduced epidemic variables into the variance equation to investigate whether the epidemic causes market fluctuations. The information has been summarized in Table 3.

**Table 3** Results of ARMA-GARCH

| Variables | S&P 500 | |
|---|---|---|
| | Closing price | Highest price |
| Mean Equation | | |
| AR | -0.52** | -0.08 |
| | (0.23) | (0.90) |
| MA | 0.38 | 0.15 |
| | (0.25) | (0.90) |
| Constant | 0.00*** | 0.00*** |
| | (0.00) | (0.00) |
| Variance Equation | | |
| New Cases | 0.04 | -0.08*** |

| | | |
|---|---|---|
| | (0.03) | (0.02) |
| ARCH | 0.36*** | 0.23*** |
| | (0.04) | (0.03) |
| GARCH | 0.63*** | 0.76*** |
| | (0.03) | (0.02) |
| Constant | -12.21*** | -12.13*** |
| | (0.43) | (0.17) |

Notes: The figures in brackets are standard errors. * refers to the significance level:

* significant at 0.1 level

** moderately significant at 0.05 level

*** highly significant at 0.01 level

### *3.2. Discussion*

We had mentioned before that it was necessary to verify whether GARCH (or ARCH) could be used in our study. From the table, we found that the coefficients of both ARCH and GARCH were highly significant, indicating that the logarithm return series of S&P 500 had obvious ARCH effect and could be modeled by GARCH.

After introducing the epidemic variable (i.e. the number of newly confirmed cases in the U.S.) into the variance equation, the coefficient of newly confirmed cases is not significant from the perspective of the closing price return rate series. This indicated that in the long run, COVID-19 did not cause volatility in the U.S. stock market. In contrast, the coefficient of newly confirmed cases was highly significant when viewed from the highest estimate. In other words, it reduced some of the volatility of the highest price return rate series for the U.S. stock market.

Similar studies had mentioned that assets such as the S&P 500-stock index had been greatly affected in the short term during the COVID-19 pandemic. Besides, the S&P 500 index had hit the 28.67% mark. As of 22 April 2020, it was −12.20% lower than the September 2019 index value. In a word, if the virus could not be controlled in the short term, its impact on the world economy would be more damaging in the long run [9]. Some other studies had also noted that new infections reported globally and in the United States magnified financial volatility. The ongoing COVID-19 crisis and related uncertainties had exacerbated volatility in US financial markets [10]. We believe that these do not conflict with our findings, as their studies examined the impact of COVID-19 on the global economy over a short period from September 2019 to January 2021 or even earlier while our study analyzed the impact on the U.S. stock market in a longer term. In addition, the United States had enacted some policies to deal with the epidemic recently, so this would also have a certain impact on the stock market. To sum up, research on COVID-19 has imitations, given that it remains a pandemic worldwide. Based on the results of our present study, we discussed that in the long term, the impact of COVID-19 on the U.S. stock market was not obvious, but it did reduce the volatility of its highest price return rate series.

## 4. CONCLUSION

In order to study the effects of major unexpected risks on the volatility of the U.S. stock market, we took COVID-19 epidemic as an example and used the Autoregressive Integrated Moving Average Model as a basis. By constructing a Generalized Autoregressive Conditional Heteroskedasticity Model, we analyzed the return series of the closing and highest prices in S&P 500 from November 4th, 2019 to August 31st, 2021. We found no significant impact on the closing price, meaning that COVID-19 does not cause volatility in the U.S. stock market over the long term. However, from the estimation results of the highest price, the new diagnosis had reduced the volatility of the return series of the highest price of the stock market. The highest and closing prices in the stock market are directly related to the trend and fluctuations of the American stock market. Through modeling and analysis of the return series of the closing and highest prices, the number of newly diagnosed patients in the United States is introduced as a variable in order to find out its impact on the U.S. stock market and the connection between the U.S. stock market and COVID-19. People who are interested in investing in the stock market but lack financial knowledge may want to read this paper so that they can reduce their losses in the event of sudden risk shocks. However, due to the pandemic trend, this study will have some limitations as the pandemic continues. To make the stock market generally stable, we may need to pay more attention to COVID-19 in the future to overcome the epidemic.

## REFERENCES

[1] Congquan Xu, Xingqiao Zhang, Kun Ding, Southwest Forestry University, 2017. Analysis of the impact of emergencies on stock market fluctuations - Evidence from Brexit.

[2] Hee Soo Lee, School of Business, Sejong University, 2020. Exploring the Initial Impact of COVID-19 Sentiment on US Stock Market Using Big Data.

[3] Xia Li, Nanjing University of Science and Technology, 2013. Research on regional electricity consumption forecasting method based on ARIMA model and regression analysis.

[4] Hou Lu, Jinan University, 2009. Short-term analysis and prediction of oil price based on ARIMA Model.

[5] Domenico Benvenuto, Marta Giovanetti, Lazzaro Vassallo, Silvia Angeletti, Massimo Ciccozzi, 2019. Application of the ARIMA model on the COVID-2019 epidemic dataset.

[6] Patrick F. Dunn and Michael P. Davis, 2017. Measurement and Data Analysis for Engineering and Science. CRC press, 479-482.

[7] Ismail E. Mohamed, 2008. Time Series Analysis Using SAS.

[8] Lamia Benzid, Sousse University, Kaouther Chebbi, King Fasail University, 2020. Impact of Covid-19 Virus on Exchange Rate Volatility: Evidence Through GARCH Model.

[9] Parag Verma, Ankur Dumka, Anuj Bhardwaj, Alaknanda Ashok, Mukesh Chandra Kestwal, Praveen Kumar, 2020. A Statistical Analysis of Impact of COVID19 on the Global Economy.

[10] Claudiu Tiberiu Albulescu, Management Department, Politehnica University of Timisoara, 2021. COVID-19 and the United States financial markets' volatility.