Volatility of the Dow Jones Pharmaceuticals and Biotechnology Index in the context of the Coronavirus crisis

Flavius DARIE\textsuperscript{a,}*\textsuperscript{,} Ileana Tache\textsuperscript{b}
\textsuperscript{a}Lucian Blaga University of Sibiu
\textsuperscript{b}Transilvania University of Brasov

\begin{abstract}
This paper’s analysis was triggered by the outbreak of the new virus COVID-19. In December 2019, the Chinese officials alerted the World Health Organization (WHO) of the existence of an unknown deadly virus. Coronavirus has rapidly spread across the world - to Europe, Middle East and the USA, forcing the World Health Organization to declare COVID-19 a global pandemic. Its spread has generated major concerns for the health and economic sectors. Meanwhile, all countries hope for the development of a vaccine. Using as a research method the EGARCH model, this paper investigates if it can be applied to model the trend of volatility of the pharmaceuticals and biotechnology markets, especially during the health crisis. More specifically, this paper tries to identify whether different specifications of univariate GARCH models can usefully anticipate volatility in the stock indices market. The study uses estimates from both a symmetric and an asymmetric GARCH models, namely GARCH (1, 1) and EGARCH models, for the Dow Jones Pharmaceuticals and Biotechnology index (DJUSPN). The dataset is extracted from “Investing.com” and covers the period September 2019 - August 2020, resulting in a total of approximately 252 daily closing prices. The data focuses on the response of the highest capitalized pharmaceutical and biotechnology companies from the US to combat the outbreak of the coronavirus. This study concludes that the EGARCH model is better than the unconditional volatility and the conditional GARCH (1, 1) volatility and it is best suited for modelling and forecasting the fluctuations of the stock indexes.
\end{abstract}

\begin{articleinfo}
\textbf{Keywords:} EGARCH model, volatility, autocorrelation, health crisis, coronavirus, Dow Jones Pharmaceutical and Biotechnology Index

*\textbf{Corresponding author:} Cosmin.darie@ulbsibiu.ro
Flavius-Cosmin Darie

\textbf{Article history:}
Received 20-10-2020
Revised 28-10-2020
Accepted 16 11-2020

1. INTRODUCTION

Volatility forecasting is vital in assessing the riskiness of an investment within the financial markets as its price significantly fluctuates according to the evaluation of volatility. The measurement of volatility becomes also critical in portfolio management. Therefore, portfolio managers and option traders find the assessment of volatility extremely valuable as it assists them in making profits while minimizing the risk of an investment. Volatility always exists in financial time series. However, there are periods with
high volatility and low volatility. It has been observed that volatility tends to significantly rise during critical economic events such as financial crisis, recessions and oil crisis (Cheteni, 2016).

According to Poon (2005), volatility is described as the spread of the likely result of an uncertain variable. Essentially, volatility is associated with risk of the investment. Levine (1991) detailed that economic development could be achieved through the elimination of firm’s impulsive liquidation of capital. Chan et al., (1997) identified that over decades, the stock markets have been strongly associated with the national economies through foreign direct investments, capital flows, international trade and technological progress.

Black (1976) laid down the theoretical grounding of the volatility of stock returns which states that the leverage effect is the main cause for stock return volatility. In other words, an increase in a company’s equity holding lowers its debt/equity ratio. The same rationality applies if a company’s equity holding falls and its debt/equity rises. Thus, the inverse relationship between the two indicators becomes evident. However, the absence of conclusiveness in the returns of the stock market has led to the foundation of numerous models evaluating the leverage effect such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models.

Raju and Ghosh (2004) identified that the stock market volatility represents a cause of interest to policy makers as the stock market uncertainty influences growth prospects and acts as a barrier for investments. For instance, the 1997 stock market crash in the United States led to a decrease in consumer spending. Therefore, volatility is regarded as an impediment through its effects on business and consumer expenditure. As a result, constant volatility shows underlying economic difficulties. The more variable the prices are, the worse and less efficient would be the pricing of securities leading to an inefficient distribution of resources.

Financial stock markets are the key drivers to development and most importantly, economic growth. Therefore, over the last three decades, modelling and forecasting stock markets volatility has become a significant setback and has drew attention to the practitioners and academics. Financial data has indicated that the conditional distribution of high-frequency returns encompasses numerous features including temporal persistence, excess kurtosis and negative skewness in conditional movements. In order to accommodate these issues, academics and econometricians have developed modelling and forecasting volatility models.

Crucial contributions to advance financial econometrics were made in 1982 by Engle with the introduction of Autoregressive Conditional Heteroskedasticity (ARCH) model and in 1986 by Bollerslev with the Generalized ARCH (GARCH). These two models became very popular due to their
ability to efficiently capture volatility clustering. According to Frances and Van Dijk (1996), the GARCH models eliminate the setback of long-term time varying volatility. A number of research papers attempted to test whether the efficient market hypothesis holds. More specifically, several researchers assessed whether excessive stock price fluctuations are due to changes in fundamentals or other factors such as incomplete information and speculative bubbles. However, one important research paper to note was the paper of Chen, Roll, and Ross (1986) which identified that stock returns are determined by economic factors and much less than changes to stock fundamentals.

The coronavirus was firstly discovered in Wuhan, where numerous cases of unusual pneumonia appeared. On 31st December 2019, the Chinese officials alerted the World Health Organization (WHO) of the existence of an unknown deadly virus. Coronavirus has rapidly spread across the world - to Europe, Middle East and the United States of America, forcing the World Health Organization to declare COVID-19 a global pandemic (CNBC, 2020).

Its spread has generated major concerns for the health and economic sectors and has left businesses and several industries worldwide in collapse. Thus, since the announcement of the new virus by the Chinese officials, the Nikkei, FTSE and Dow Jones Industrial Average have all massively plummeted. Similarly, the oil price is at its lowest level since 2001 and even the value of a safer investment, gold, has greatly decreased. Furthermore, due to travel restrictions, the tourism industry has enormously suffered with airlines, hotels and restaurants taking the biggest hit. Each country around the world has taken distinctive measures in an attempt to protect their national economy during the health crisis (BBC, 2020).

Meanwhile, all the countries hope for the development of a vaccine against the new virus. Thus, the main focus has now been put on the drug companies since heavy investments were made, especially by the US government, in order to combat the virus. Thus, the President of the United States hosted executives from Pfizer, Johnson & Johnson and several other drug companies to discuss efforts to create a vaccine and therapeutics against the new deadly virus (CNBC, 2020).

As of July 2020, the US economy suffered a bit hit and appears to be in deep trouble due to the coronavirus. Restaurants reservations are decreasing, the foot traffic at stores is dropping and the rebound of the air travel seems to have stabilized to the new reduced demand. Therefore, there is mounting evidence that the economic recovery in the United States is already stalling as the number of deaths and infections due to the coronavirus is dramatically rising. Unemployment claims remain elevated. Real-time economic indicators bottomed out in May 2020 as stay-at-home directives were lifted and many American citizens felt safe enough to return to their normal routine such as visiting restaurants, shopping centers and even airports. That gave hope prematurely as the number of coronavirus cases rapidly increased and caused the growth in the economy to slow (BBC, 2020).
However, on the vaccine front, real progress is being made highlighted by a $1.95 billion deal for Pfizer to produce millions of Covid-19 vaccine doses for the US government (CNN Business, 2020).

This paper analyses the fluctuations of the Dow Jones Pharmaceuticals and Biotechnology index (DJUSPN) from September 2019 to August 2020. The study uses estimates from both a symmetric and an asymmetric GARCH models, namely GARCH (1, 1) and EGARCH models, for the Dow Jones Pharmaceuticals and Biotechnology index in order to see the behaviour of volatility of the mentioned index during the major pandemic across the United States. In other words, the objective is to verify to what extent the trend calculated by the GARCH (1, 1) and EGARCH model matches with the events during the mentioned period and which of the two models provides the best fit to the chosen sample.

2. METHODOLOGY AND DATA

In this section the GARCH models are described both from a statistical and financial perspective, based on Tache and Darie (2019a and 2019b). Both the leverage effect and volatility clustering are components of the GARCH methodology by adding to the model of linear regression the conditional equation. Below, we present both the symmetric and the asymmetric GARCH models that help to investigate the main features of volatility. As stressed in our preceding papers (Tache and Darie, 2019a and 2019b), the estimation using the maximum likelihood approach highlights the stability of the GARCH models, the choice of data period and the way it affects long-term volatility.

As Alexander (2001) points out, in a generalized autoregressive conditional heteroscedasticity model, returns are considered to be generated by a stochastic process with volatility depending on the time interval; instead of modelling the data after having been collapsed in a single unconditional distribution, a GARCH model takes into account more detailed assumptions regarding the conditional distributions. As conditional variance is an autoregressive process, the conditional distributions change in time in an auto-correlated way.

Our research uses both the symmetric and the asymmetric GARCH models, namely the GARCH (1, 1) and the EGARCH models. This paper presents the advantages of the EGARCH model as compared with the symmetric GARCH (1, 1) model: a) allowing asymmetries for capturing the leverage effect and b) using log returns for obtaining a positive conditional variance, even with negative parameters. Unlike the symmetric GARCH model in which a symmetric response of volatility to both positive and negative shocks will be produced when a shock appears, the asymmetric GARCH models involve asymmetric responses, illustrating that positive shocks will engage a lower volatility than negative ones. Among the various mathematical alternatives of the GARCH (1, 1) and the EGARCH models, we decided to use Alexander’s version (2001).
This paper works with the following exponential GARCH equation:

\[
\ln(\sigma_t^2) = \omega + g(z_{t-1}) + \beta \ln \sigma_{t-1}
\]

\[
g(z_t) = \theta z_t + \gamma \left( |z_t| - \frac{2}{\pi} \right)
\]

\[
z_t = \frac{u_t}{\sigma_t}
\]

Where \( \sigma_t \) represents the conditional variance as an asymmetric function of lagged disturbances; \( u_{t-1} \), \( g(z_t) \) is a linear asymmetric response function in \( z_t \) with slope coefficient \( \theta + 1 \) in case \( z_t \) is positive while \( g(z_t) \) is linear with \( z_t \) with slope coefficient \( \theta - 1 \) in case \( z_t \) is negative. As a result, large innovations increase the conditional variance in case \( |z_t| - E|z_t| > 0 \) and decrease the conditional variance in case \( |z_t| - E|z_t| < 0 \) only while \( \theta = 0 \). On the other hand, the innovation in variance \( g(z_t) \) is positive in case the innovations \( z_t \) are less than \( \frac{\sqrt{2}}{\theta - 1} \). Thus, the negative returns \( u_{t-1} \) cause the innovation to the conditional variance to be positive in case \( \theta \) is much less than 1 (Tache and Darie, 2019a and b).

This paper works with the following GARCH \((1, 1)\) equation:

\[
\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2
\]

\[
\omega > 0; \ \alpha, \beta \geq 0
\]

Where \( \omega \) is the constant, \( \alpha \) is the GARCH error coefficient, \( \beta \) is the GARCH lag coefficient and \( \sigma_t^2 \) is the conditional variance since any past information considered to be relevant is included in the one period ahead estimation of calculated variance. While the unconditional variance of the GARCH model is constant and concerned with long-term behaviour of time series, the conditional variance relies on the past information (Tache and Darie, 2019a and b).

The formula for unconditional variance of the GARCH \((1, 1)\) model is presented below:

\[
Var(u_t) = \frac{\omega}{1-(\alpha_1+\beta)}
\]

The coefficient measures the degree to which today’s volatility shock is encompassed within the volatility of the next period; in other words, it relates to the long-term volatility. The unconditional variance remains constant for as long as \( \alpha_1 + \beta \) is strictly lower than 1 (Tache and Darie, 2019a and b).
As Alexander (2001) explains, GARCH models are often estimated on intraday and daily data in order to capture volatility clustering effects in the returns of financial assets. These effects disappear with returns observed during long term periods. The GARCH parameters will be estimated by maximizing the value of the log likelihood function – using time varying mean and variance.

The following maximizing problem is to be solved when trying to maximize the GARCH (1, 1) likelihood:

$$\ln L(\theta) = -\frac{1}{2} \sum_{t=1}^{T} (\ln (\sigma_t^2) + \frac{(\varepsilon_t^2)}{\sigma_t^2})$$

Few convergence problems will be met when maximizing the log likelihood function for univariate GARCH models. Changes of data may induce changes in the coefficient estimates, but if there are not real structural breaks in the data generation process, the estimates of the parameter do not significantly change with the new data introduction.

A proper definition of the log likelihood function involves a certain minimum quantity of data. Sometimes, an adequate convergence of the model needs many years of daily data. However, the data within this study covers only a short period of time, namely from September 2019 to August 2020, resulting in a total of 252 daily observations. This period was specifically chosen to reflect the behaviour of volatility before and during the coronavirus pandemic.

3. EMPIRICAL RESULTS AND ANALYSIS

The early evidence for supporting the use of the ARCH/GARCH models is autocorrelation. In order to identify the existence of autocorrelation within the dataset, we used the Box-Pierce or Q test, according to Alexander (2001). The Box-Pierce test helps to determine whether a time series consists simply of random values (white noise). This test was made for the residuals of the time series, after fitting to the data an ARCH (p, q) model. The formula presented below was used for identifying the autocorrelation:

$$Q = n \sum_{k=1}^{\min(m,h)} r_k^2$$

Where Q represents the Box-Pierce statistic, n represents the total number of observations, m is the number of parameters and h represents the maximum lag considered.

The Box-Pierce test indicates either H0: Prices do not have any significant historic dependence, or H1: Prices do have significant historic dependence. In general, the Q test statistic shows that if residuals are
white noise, the Q statistic will have a $\chi^2$ distribution with (h-m) degrees of freedom. If each rk value is close to zero, the Q statistic will be very low; but if some rk values are large, the Q test statistic will be high. So, a comparison between the Q statistic with $\chi^2$ distribution becomes necessary.

![Diagram of autocorrelation](image)

Figure 1. DJUSPN Autocorrelation for the period September 2019 to August 2020 (elaborated by the authors)

Figure 1 presented above shows that autocorrelation in returns does exist for the Dow Jones Pharmaceuticals and Biotechnology index for the period September 2019 to August 2020. More specifically, figure 1 displays the autocorrelation line within the calculated confidence interval of 95%. Therefore, the confidence interval can be interpreted as the range within which the true value is likely to lie with 95% confidence.

However, in order to test for autocorrelation, the application of the Q test or Box-Pierce test is also made. More specifically, only the first six lags will be highlighted within this paper and therefore, focus on the rk values for only the first six observations. Accordingly, the Box-Pierce process for the DJUSPN Index between September 2019 and August 2020 is presented below:

$$Q = 252 \sum_{k=1}^{6} r_k^2 = 70.15$$

As presented above, the Box-Pierce test for the specified time period is paralleled with the critical value of the Chi-squared of 12.5 for a 5% significance level. Since the Box-Pierce test result shows a value higher than 12.5, $H_1$ is the accepted hypothesis as returns do have significant historic dependence.
Table 1 and Table 2 illustrate the manner in which the maximum likelihood function was estimated using Excel Solver.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DJUSPN</strong></td>
<td><strong>GARCH (1, 1)</strong></td>
</tr>
<tr>
<td>( \omega )</td>
<td>1.21E-04</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.3545</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.6455</td>
</tr>
<tr>
<td>( \alpha + \beta )</td>
<td>1.0000</td>
</tr>
<tr>
<td>LT Vol</td>
<td>12073664.02%</td>
</tr>
<tr>
<td>LogL</td>
<td>941.1996</td>
</tr>
<tr>
<td>Unconditional Vol</td>
<td>27.57%</td>
</tr>
</tbody>
</table>

More precisely, table 1 and table 2 show the estimated values of the parameters presented in the GARCH (1,1) and EGARCH models that contribute to the calculation of the maximum value of the log likelihood and the long-term volatility for the DJUSPN Index between September 2019 and August 2020. Note the difference between the traditional model which uses unconditional volatility estimated by equally weighted average of all squared returns of 26.90% and the long-term volatility of 16.90% provided by the EGARCH approach. On the other hand, the GARCH (1, 1) approach provides a massive value for the long-term volatility of 12073664.02% as any past information is reflected in the one period ahead estimation of calculated variance. In other words, GARCH (1, 1) estimated volatility recorded substantial spikes of approximately 140% compared to the EGARCH estimated volatility of approximately 80%. Due to the assimilated massive spikes during the mentioned period, the long-term volatility provided by the GARCH (1, 1) model records such a high percentage.

All the parameters encompassed within the EGARCH and the GARCH (1, 1) models are optimally calculated utilizing the maximum likelihood approach. Due to the COVID-19 pandemic, the value of the coefficient \( \beta \) indicates that there is a high persistence in volatility for the EGARCH model whereas the coefficient \( \beta \) shows only a moderate persistence in volatility for the GARCH (1, 1) model. Contrary to the traditional approach where the volatility forecasts are constant, the GARCH (1, 1) and the EGARCH models provide volatility forecasts with the estimated parameters changing as new data arrives. Furthermore, both the GARCH (1, 1) and the EGARCH approaches offers a higher intensity when any major event is occurring, but the term structure of the estimated volatility will mean-revert to the long-term average (Alexander, 2001).
According to Alexander (2001), the parameters encompassed in the both of the models are very sensitive to the dataset particularly the constant, $\omega$. On this occasion, the dataset covers nearly one year during which a major health crisis occurred in not only the United States but worldwide. The estimate of the constant ($\omega$) is not extremely high for both of the models to show any disturbances in the dataset that was used. Also, the long-term volatility provided by the EGARCH calculation appears to be far lower than the value of the unconditional volatility and the GARCH (1, 1) estimated volatility. As stated above, the reason for such a major discrepancy in the value of long-term volatility and unconditional volatility is the fact that the long-term volatility provided by the GARCH models is not constant over a period of time and shows the unexpected market shock at a higher intensity until the effects produced by the market shock disappears. Moreover, any major shock regardless the industry could threaten any stable market. In order for the estimated parameters to be stable, there is a trade-off between having enough data and too much data so that the long-term volatility provided by the GARCH estimates to reflect as good as possible.

The $\gamma$ parameter measures the asymmetry or the leverage effect. Such a parameter is vital to the analysts who use the EGARCH model. Since $\gamma$ shows a value greater than 0, more specifically 0.09610, then positive shocks (good news) generate much more volatility than negative shocks (bad news) (Brooks, 2014). Therefore, any good news regarding a vaccine or treatment from the US Pharmaceutical and Biotechnology companies for the COVID-19 pandemic would make a greater impact than any other negative news.

![Figure 2. Comparison between EGARCH Estimated Volatility versus Unconditional Volatility for the DJUSPN Index from September 2019 to August 2020 (elaborated by the authors)](image-url)

Figure 2 presents that the i.i.d unconditional volatility estimate of 26.90% is higher than the unconditional EGARCH volatility of 16.90% whereas figure 3 presents that the conditional GARCH (1, 1) estimated volatility of 12073664.02% is much higher than the i.i.d unconditional volatility.
Both of the presented GARCH models efficiently capture the trend of volatility. However, the conditional GARCH (1, 1) approach overestimates the negative shocks compared to the unconditional EGARCH approach which provides much more reliable forecasts as it efficiently captures the negative shocks. However, it is not unusual to find differences between the i.i.d volatility and the long-term GARCH volatility models since both of the models does not assume that returns are i.i.d. Furthermore, until February 2019, the fluctuations in the volatility of Dow Jones Pharmaceutical and Biotechnology Index does not raise concerns regarding any uncommon behaviour. However, in March 2019, the volatility of the Dow Jones Pharmaceutical and Biotechnology Index spikes dramatically as the World Health Organization declares the outbreak of coronavirus a pandemic. Furthermore, the President of the United States declares national emergency to unlock $50 billion in federal resources to combat the coronavirus and signs into law a coronavirus relief package (CNN, 2020).

As of April 2020, the volatility of the Dow Jones Pharmaceutical and Biotechnology Index gradually decreases due to safety economic measures taken by the US federal government to provide funding for some of the workforce obliged to stay-at-home. Furthermore, during May 2020, the volatility of the DJUSPN Index seems to decrease even more as the US and his administration expands the coronavirus testing capabilities by sending $11 billion to states. However, a much lower spike in the volatility of the DJUSPN Index is recorded as the administration of the United States notifies Congress and the UN that the US if formally withdrawing from the World Health Organization (CNN, 2020). Until 31st August,
there has been no other major spike in the volatility of the DJUSPN Index as a vaccine is being developed highlighted by a $1.9 billion deal for Pfizer to produce millions of Covid-19 vaccine dozes for the US government (CNN Business, 2020). As of 31st August, the US holds the first position regarding the total number of deaths worldwide (The Guardian, 2020). Therefore, having signed a deal for the vaccine is crucial for the American population as it provides some relief during this health crisis. It must be mentioned that Pfizer is part of the Dow Jones Pharmaceutical and Biotechnology Index.

4. CONCLUSION

After the World Health Organization declared a global coronavirus pandemic, the volatility of the Dow Jones Pharmaceutical and Biotechnology Index drastically increased as the US holds the first position regarding the total number of deaths worldwide. However, the coronavirus pandemic in the US has been taken very seriously with the US government having prepared relief packages for the working force and substantial stimulus for the US economy. Furthermore, the signing of a deal for the coronavirus vaccine offers some relief for the US citizens. Since the events presented in this paper matched with the volatility trend provided by the EGARCH model, it must be mentioned that the GARCH (1, 1) and the EGARCH models accurately capture the volatility trend of the Dow Jones Pharmaceutical and Biotechnology. Furthermore, this paper highlights the significant difference between using the volatility provided by the GARCH (1, 1) and the EGARCH model against the unconditional volatility. However, the conditional GARCH (1, 1) approach overestimates the negative shocks as compared with the unconditional EGARCH, the latter providing much more reliable forecasts as it efficiently captures the negative shocks. In conclusion, having as a case study the health industry, the EGARCH model represents the best model of the three models which forecast the volatility of the DJUSPN index as it precisely captures the trend of the volatility without overestimating the results.

CONFLICT OF INTEREST AND PLAGIARISM: The authors declare no conflict of interest and plagiarism.

REFERENCES


