# FORECASTING VIX INDEX AS A MEASURE OF MARKET VOLATILITY BY THE USE OF GOOGLE QUERIES

**Bella GRIGORYAN**
PhD student in Mathematical Economics, YSU
**Tigran GRIGORYAN**
MS student in Electronics, NPUA

Key words: Google Trends; Google searches; VIX Index; ARDL; volatility

***Introduction.*** Understanding and predicting human behavior is of paramount importance. As the Internet penetration is constantly growing, users' Web-based search patterns are valuable for facilitating a better understanding of human behavior [1]. The study conducted by Jun et al. reveals that Google Trends, which has dramatically increased in popularity among researchers in the last decade, is used to analyze and forecast various variables in different areas, including IT, health, business and economics [2].

Traditionally, there has been a plethora of evidence that market agents are rather sensitive to the oil price changes. Rocketing or plunging of oil prices results certain volatility response in the market. The study carried out by Bastianin A., Manera M. suggests volatility responds significantly to oil price shocks caused by unexpected changes in aggregate and oil-specific demand, while the impact of supply-side shocks is negligible [3]. As per another study conducted by the same authors [4], the stock market volatility is not significantly affected by oil supply shocks, whereas it is very sensitive to oil demand shocks.

***Scientific novelty***. In this paper, we show that Google searches can help to forecast the CBOE Volatility (VIX) Index [5], a popular measure of the market's expected volatility on the S&P 500 Index [6], calculated and published by the Chicago Board Options Exchange (CBOE). VIX Index is considered as a leading indicator and should not be thought of as an immediate S&P 500 Index movement. Peter Carr substantiates that VIX can be considered as a fear gauge [7]. VIX Index is known to capture the market sentiment. In other words, it generally depicts the level of investor anxiety. In this paper we investigate if Google searches can also reveal investor anxiety, thus making Google queries a useful tool to understand the volatility movements. We hypothesize that oil-related Google queries, as a proxy of volatility expectations and/or anxiety of market agents, can be used to forecast VIX values. To the best of our knowledge, no study has been conducted to forecast VIX index by the use of Google queries related to oil so far. In this paper, we show the relationship between Google searches on oil and related terms and VIX Index. Furthermore, we try to forecast VIX Index based on the model under the study.

***Literature Review.*** As the popularity of Google Trends is growing, scientific society are trying to utilize the data it provides to get more information on the financial markets. A study conducted by Callet, D. in 2013 [8] concludes that it is highly debatable if Google Trends contain enough information that is useful to predict financial markets. Contrary to this, in their paper Preis et al. [9] suggest that Google Trends contain information reflecting future trends of human behavior, thus it can have predictive power from financial markets' perspective. Using the historical data from 2004-2011, it is shown that before stock markets start falling usually there is an increase in Google search volumes for specific keywords relating to financial markets. This means that these warning signs in search volume data could have been exploited in the construction of profitable trading strategies.

Another study carried out by Ahmed et al. [10] claims a strong correlation between financial decision making and human behavior. To be more specific, the authors analyze and predict Karachi Stock Exchange 100 index considering the impact of the political and business events, and thus considering the corresponding Google searches. The results of the study indicate that the change in the searches of particular topics on Google may lead to stock market fall or rise. A study conducted by Habibah et al. [11] tries to answer several questions. First of all, it shows that there is a significant positive correlation between VIX and Google Indices as both of them can be considered as "sentiment" indices. Secondly, it compares both VIX and Google indices to check which one of the indices better captures the market pessimistic sentiments. Results suggest that change in VIX contains information that helps to forecast the Market Crash and Bear Market, and Google sentiment indices (Market crash and Bear market) also contain some information to explain the change in VIX.

***Methodology.*** Autoregressive distributed lag (ARDL) modelling was applied to investigate the relationship between VIX index and oil queries on Google. To that end, weekly VIX Index (^VIX) values for the time period of 23.10.2016 - 27.12.2020 and the corresponding weekly oil searches on Google (search for the terms oil price, price of oil, and crude oil). It is worth noting that Google trends does not provide us with the absolute number of the queries, but rather with the relative numbers compared to the peak popularity of the term, which is the maximum number of searches for the given period, which is considered to be 100. We included the percentage change of relative oil searches in the model, which is equivalent to including the percentage change of oil searches in absolute numbers. Stationarity was tested with Augmented Dickey-Fuller Unit Root test [12,13]. The following candidate models were compared: all ARDL models with up to 3 lags for both regressor and the dependent variable, and also the model obtained by removing statistically non-significant lags from the best ARDL model. Akaike Information Criterion (AIC) [14] was used for model selection. Additionally, out-of-sample static forecast was made with the best model for the period of 27.12.2020 – 17.10.2021.All inference was made at 5% significance level. Statistical computing was done with EViews 9.0 statistical package.

***Analysis.*** *Model Estimation Results*. Table 1 illustrates that for both series we can reject the null hypothesis of having a unit root at 5% significance level.

**Table 1**. Augmented Dickey-Fuller test results

| | p-value* |
|---|---|
| VIX | 0.004 |
| Oil searches: percentage change | <0.001 |
| *p-values were obtained from Augmented Dickey-Fuller test with maximum of 15 lags.* | |

The following candidate models were estimated:

Model 1: $VIX_t = 2.4 + 0.79VIX_{t-1} + 0.09VIX_{t-2} - 0.03VIX_{t-3} + 1.4OIL_t + 3.29OIL_{t-1} + 4.71OIL_{t-2} - 0.55OIL_{t-3}$

Model 2: $VIX_t = 2.31 + 0.86VIX_{t-1} + 1.2OIL_t + 3.23OIL_{t-1} + 4.44OIL_{t-2}$

Model 3: $VIX_t = 2.22 + 0.87VIX_{t-1} + 2.97OIL_{t-1} + 4.28OIL_{t-2}$

Model estimation results for all candidate models are shown in Table 2 below:

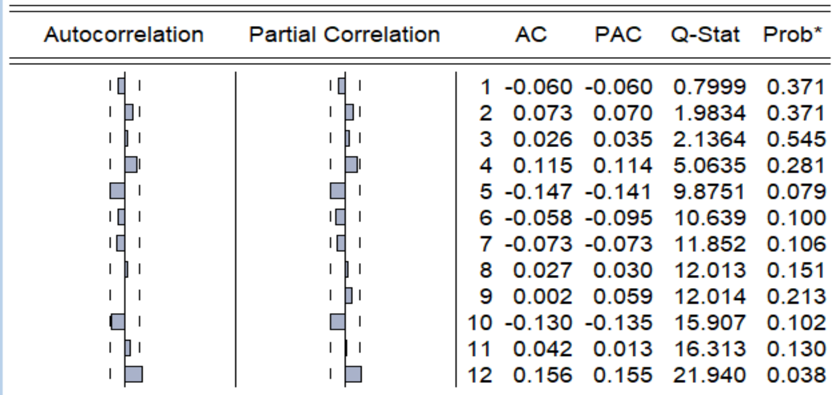
**Table 2.** Model Estimation Output.

| **Variable** | **Model 1** | **Model 2** | **Model 3** |
|---|---|---|---|
| | **p-value** | | |
| $VIX_{t-1}$ | <0.001*** | <0.001*** | <0.001*** |
| $VIX_{t-2}$ | 0.335 | - | - |
| $VIX_{t-3}$ | 0.673 | - | - |
| $OIL_t$ | 0.079 | 0.121 | - |
| $OIL_{t-1}$ | 0.518 | <0.001*** | <0.001*** |
| $OIL_{t-2}$ | <0.001*** | <0.001*** | <0.001*** |
| $OIL_{t-3}$ | <0.001*** | - | - |
| | **AIC [1]** | | |
| | 5.4803 | 5.4628 | 5.4649 |
| [1] Akaike Information Criteria. | | | |

As it can be seen from Table 2, the following model has the lowest AIC among all candidate models:

$$VIX_t = 2.31 + 0.86VIX_{t-1} + 1.2OIL_t + 3.23OIL_{t-1} + 4.44OIL_{t-2}$$

We can infer from the equation that, on average, increase in the percentage change of oil searches in the given week and in up to the previous 2 weeks is associated the increase in VIX index. It is also worth displaying the autocorrelations and partial autocorrelations of the residuals for the selected model. Figure 1 below illustrates that the autocorrelation and partial autocorrelation coefficients up to the 12$^{th}$ lag are not significantly different from 0, thus residuals can be considered as serially non-correlated.



| | AC | PAC | Q-Stat | Prob* |
|---|---|---|---|---|
| 1 | -0.060 | -0.060 | 0.7999 | 0.371 |
| 2 | 0.073 | 0.070 | 1.9834 | 0.371 |
| 3 | 0.026 | 0.035 | 2.1364 | 0.545 |
| 4 | 0.115 | 0.114 | 5.0635 | 0.281 |
| 5 | -0.147 | -0.141 | 9.8751 | 0.079 |
| 6 | -0.058 | -0.095 | 10.639 | 0.100 |
| 7 | -0.073 | -0.073 | 11.852 | 0.106 |
| 8 | 0.027 | 0.030 | 12.013 | 0.151 |
| 9 | 0.002 | 0.059 | 12.014 | 0.213 |
| 10 | -0.130 | -0.135 | 15.907 | 0.102 |
| 11 | 0.042 | 0.013 | 16.313 | 0.130 |
| 12 | 0.156 | 0.155 | 21.940 | 0.038 |



**Figure 1.** Correlogram of the residuals: Model with the lowest AIC.

* Probabilities may not be valid for the equation specification.

Note: It can be seen that none of the correlation coefficients (up to the 12th lag) is significantly outside the marked boundaries, which means none of the coefficients is significantly greater than 0 in absolute value.

*Model Forecasting*. The selected model was used to make an out-of-sample forecast for 27.12.2020 – 17.10.2021. Figure 2 illustrates the forecasted and actual graphs for VIX Index. In addition, Figure 3 illustrates the actual VIX values along with the forecasted values ± standard deviation.

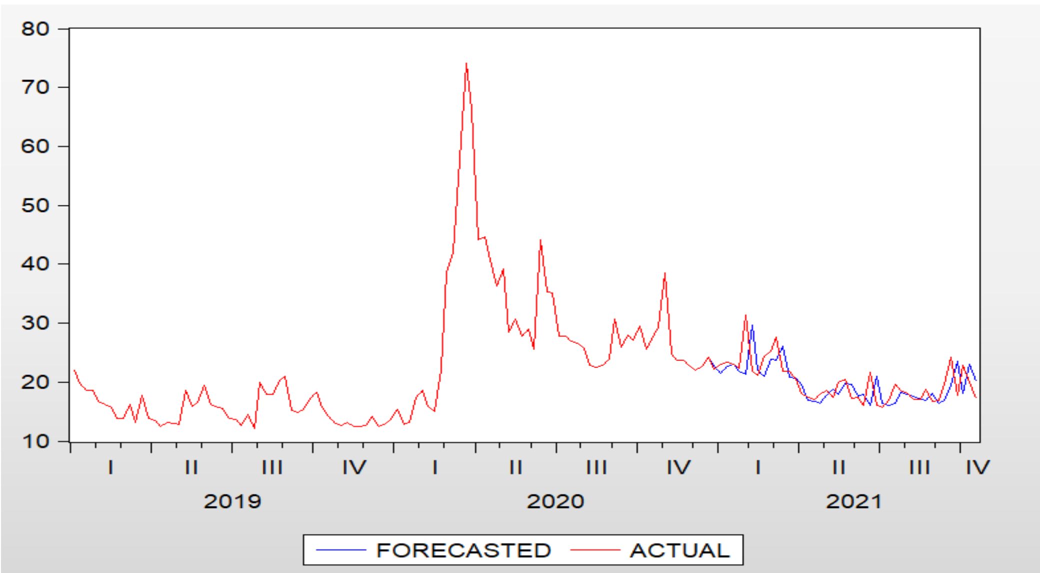



**Figure 2.** Actual and Forecasted VIX Values.

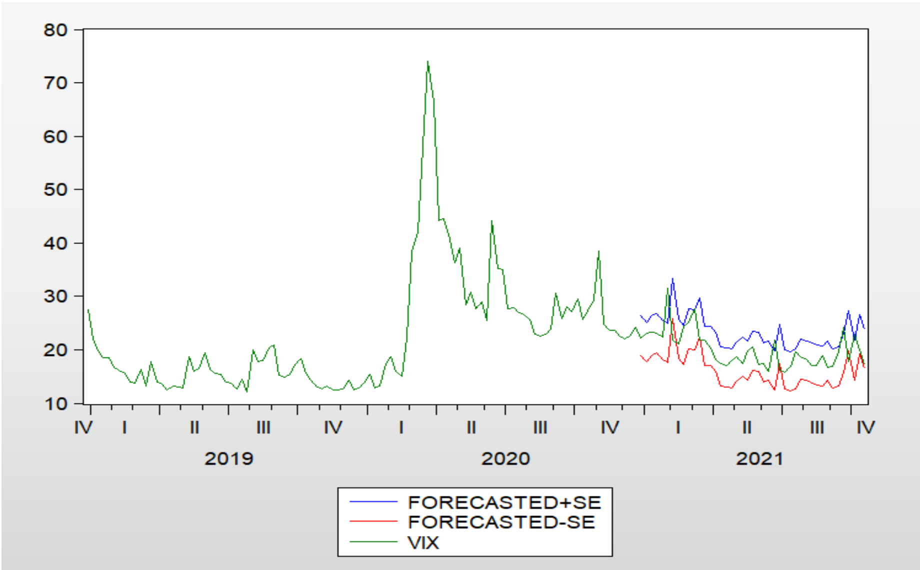



**Figure 3.** Actual VIX values with forecasted values +/- SE.

***Conclusion.*** Regardless of the growing popularity of Google Trends data in the recent decade, we believe that it still has enough room to grow in popularity considering the potential predicitive power the data it provides might contain. Although several authors tried to answer various questions on the usage of Google Trends data in financial markets, to the best of our knowledge, there are no previous studies carried out to forecast VIX Index using the Google searches on oil and related terms. The results of the study are not only informative from the perspective of the question under the study in this paper, but also they might provide food for thought for other researchers to further investigate the topic.

The study results indicate that there is a statistically significant relationship between Google queries on oil and market volatility. To be more precise, percentage changes in oil searches in the given week, the previous week, and 2 weeks ago are associated with the volatility change measured by VIX Index in the same direction in the given week. This seems reasonable, since both factors reflect human behavior the same way. When people have either well-grounded expectations or worries related to the macroeconomic situation, they start with searching for more information on Google to understand the situation better and act to the best of their knowledge. Oil is considered to be some kind of indicator of how the economy behaves. On the other hand, once people fear, VIX also increases. To sum up, oil searches capture ether optimistic or pessimistic sentiment in the market, which in turn

## References


1. Mavragani A.; Ochoa G.; Tsagarakis K., Assessing the Methods, Tools, and Statistical Approaches in Google Trends Research: Systematic Review. 2018, 20(11): e270, pp. 1-20.
2. Jun S., Yoo H., Choi S., 10 years of research change using Google Trends: From the perspective big data utilizations and applications. TFSC 2018, 130, pp. 69-87.
3. Bastianin, A.; Manera, M., How Does Stock Market Volatility React to Oil Price Shocks? Macroeconomic Dynamics 2017, 22, pp. 1-17.
4. Bastianin, A.; Conti F.; Manera, M., The Impacts of Oil Price Shocks on Stock Market Volatility: Evidence from the G7 Countries, Energy Policy 2017, vol. 98, issue C, pp. 160-169.
5. Chicago Board Options Exchange (CBOE), White Paper: CBOE Volatility Index, available at: https://cdn.cboe.com/resources/futures/vixwhite.pdf (last accessed on 31Oct2021).
6. S&P Dow Jones Indices, available at: https://www.spglobal.com/spdji/en/indices/equity/sp-500/#overview (last accessed on 31Oct2021).
7. Carr P., Why is VIX a fear gauge? Risk and Decision Analysis 2017, 6, pp. 1–7.
8. Challet, D.; Bel Hadj Ayed; A., Predicting Financial Markets with Google Trends and Not so Random Keywords, SSRN Electronic Journal 2013, doi:10.2139/ssrn.2310621.
9. Preis, T.; Moat, H. ; Stanley, H., Quantifying Trading Behavior in Financial Markets Using Google Trends, Sci Rep 3, 1684 (2013), https://doi.org/10.1038/srep01684
10. Farrukh A.; Dr. Raheela A.; Dr. Saman H.; Muhammad M., Financial Market Prediction using Google Trends, International Journal of Advanced Computer Science and Applications(ijacsa), 8(7), 2017, http://dx.doi.org/10.14569/IJACSA.2017.080752
11. Habibah U., Rajput S., Sadhwani R., Stock market return predictability: Google pessimistic sentiments versus fear gauge, CEF, (2017), Taylor & Francis, Abingdon, Vol. 5, Iss. 1, pp. 1-15
12. Dickey, D. A.; Fuller, W. A.,Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association 1979, 74, pp. 427-431.
13. Said, S. E; Dickey, D. A.,Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, Biometrika 1984, 71, pp. 599-607.
14. Bozdogan H., Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika 1987, 52, pp. 345–370.


**Bella GRIGORYAN, Tigran GRIGORYAN**

**Forecasting VIX index as a measure of market volatility by the use of Google queries**

*Keywords: Google Trends; Google searches; VIX Index; ARDL; volatility*

Modelling human behavior is rather challenging as imitating it with proxy variables is not straightforward. In recent years, search engines collect and provide us with a plethora of data, which might be a rather effective way of analyzing or forecasting human behavior. Although several authors tried to answer various questions on the usage of Google Trends data in financial markets, to the best of our knowledge, there are no previous studies carried out to forecast VIX Index using the Google searches on oil and related terms. In this paper we use Google searches on oil and related terms as a proxy variable for human expectations to model the CBOE Volatility Index. To that end, tradetional ARDL modelling was applied. The results indicate that there is statistically signify-cant relationship between Google queries on oil and market volatility. We explain this from the perspective of decision making since certain search activities on Google reveal the urge to show certain behavior and, on the other hand, the same behavior affects the market volatility.