Empirical analysis of security papers of high-technology companies on the basis of a VARMA model

Tatiana Kudryavtseva 1, Evgeniia Kozlova 1, Angi Skhvediani 1*

1 Peter the Great St. Petersburg Polytechnic University, Politechnicheskaya st., 29, St. Petersburg, 195251, Russia

* E-mail: shvediani_ae@spbstu.ru

Abstract. A VARMA-model represents a modern instrument for modeling and forecasting time series having a high degree of novelty for national science. An algorithm of study on the basis of a VARMA-model is suggested in the article, then it is tested through the example of modeling dynamics of price increment for security papers of Apple, Microsoft Corporation and Netflix. Within the framework of the study undertaken, the influence of historical data of trading on change in the future price of security papers of the companies being analyzed has been revealed: interconnections between shares of each company have been evaluated, and also the influence of the change of the shares dynamics of one company on quotations of two other companies has been analyzed. The developed model has been tested for robustness, also its validity has been evaluated and measures for optimization of VARMA-models have been suggested. Besides, timeliness of its use when undertaking national studies has been substantiated.

1. Introduction
Modeling and forecasting market value of shares is a topic of interest when undertaking studies in the field of stock market. In the national and foreign practice there are two main approaches to analysis of shares – technical and fundamental. When using these two approaches holistically, one can, with a high degree of probability, forecast the dynamics of price change for quoted shares, taking into account historical trade data and also the existing situation and also market activity indicators for companies [1–3].

The dynamics of quotations of Apple, Microsoft Corporation and Netflix shares, quoted in NASDAQ and included into NASDAQ-100 index, has been analyzed in the study. NASDAQ listing includes principally high-technology companies and NASDAQ-100 index includes, correspondingly, 100 high-technology companies with the highest market capitalization. Belonging to science-based economy sectors is a characteristic feature of high-technology companies. Thus, Apple is one of the leading high-tech companies in the world, Microsoft is a leader in the field of software engineering and Netflix is a leading entertainment Internet-service in the world.

Technical analysis on the basis of historical data forms the basis of analyzing quoted securities. Data from trading are used in the analysis: from the technical point of view, they represent high-frequency time series that is statistical data on dynamics of value change for a parameter or parameters in time with minimum intervals. Access to them is possible through a limit order book data (LOB).
One of such bases is LOBSTER – a data base that grants access to trading data for academic and scientific research goals [4].

High frequency of these data and also the necessity to consider interconnections between several time series makes their analysis, only with the help of the economic theory, very difficult to carry out, therefore it is reasonable to apply mathematics-statistical methods for their analysis and forecasting.

2. A VARMA-model

In the national and foreign practice, a set of econometric models is used for the analysis of time series – VAR (vector regression with a moving average), VMA (vector moving average), ARMA (autoregression with a moving average), AR (autoregression), and a modified VARMA-model as well (vector autoregression with a moving average). A VARMA-model is applied in foreign studies, however, it has a high degree of novelty in the Russian practice – it is not represented in special scientific literature [5–7].

Elaboration of the VARMA-model was based on the success of the ARMA-model suggested by Box and Jenkins in 1970es and also the VAR-model suggested by Sims in 1980. The ARMA-model has proved to be an accurate instrument for modeling and forecasting time series, however it had a disadvantage: it was possible to use only one variable in the analysis, though very often, when undertaking econometric studies, it is necessary to analyze interconnections between several variables. On the contrary, the VAR-model accounted for interconnections between variables as it is possible to use several variables when developing a model. Within this context, it was reasonable to elaborate the method of modeling and forecasting time series which allowed to use several variables in the analysis and take into account connections between them [8, 9].

The VARMA-model represents a modern econometric method with the use of which one can solve the problem of modeling several time series and forecasting the dynamics of their changes in future periods.

Mathematical notation of the VARMA-model is represented below:

$$Y_t = c + A_t Y_{t-1} + \ldots + A_{p} Y_{t-p} + \varepsilon_t + B_t Y_{t-q} + \ldots + B_{q} Y_{t-q}$$

where,

- $Y_t$ – dynamics of temporal variations of the value being analyzed;
- $\varepsilon_t$ – magnitude of error, $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_n\}$;
- $c$ – constant coefficient;
- $A_1, A_p, A_{p-1}, \ldots, A_{1-p}, B_1, B_{q}, B_{q-1}, \ldots, B_{q-q}$ – parameters of KxK matrices.

In order to develop a VARMA-model, time series, used in the analysis, should be stationary that is their probabilistic properties should not change in the course of time.

In contrast to most sciences, time series in economics have a trend – their averages change in the course of time. It has been generally accepted for a long time that such time series can’t be put in a stationary form. However, Box and Jenkins suggested a method of time series analysis based on the analysis of previous values of the time series which is, in substance, a complicated procedure of extrapolation, by suggesting a time series fixation method by means of transition to a series of differences after logarithm of the series [10]. Transition can be implemented by the equation below (2):

$$\text{dif} \ln(Y_t) = \ln(Y_t) - \ln(Y_{t-1})$$

3. Method

In the analysis carried out, we suggested the following algorithm for study on the basis of the VARMA-model:

1. forming a primary statistical information base;
2. primary analysis of original data;
3. putting statistical data into the stationary form in accordance with the requirements of the VARMA-model;
4. identification (selection of form) of the VARMA-model \((p, q)\) – selection of parameter values, \(p\) and \(q\), where \(p\) is the number of lagged values of the variable being analyzed which represents the order of the auto regression part of the model (AR), and \(q\) is the number of lagged values of an error which represents the order of the moving average part of the model (MA);
5. diagnostics of the model – testing the developed model for robustness. Robustness – a property of a statistical model which is characterized by the independence of its influence on the result of exploring different types of outliers, resistance to interference. A test for robustness means appearance of outliers, reduction of their influence on the result or excluding them from the sample.
6. estimation of fidelity of the model results. In order to estimate fidelity of the developed models, AIC and BIC criteria are applied and also graphical superposition of the simulated result upon the real result is used;
7. forecasting;
8. interpretation of the result:
9. in case of inaccurate modeling and forecasting (bad-fit) at the 6-th and 7-th stages – return to item 4;
10. in case of accurate modeling and forecasting (good-fit) – to item 9;
11. making a management decision on the basis of the analysis carried out.

Graphically, the algorithm of carrying out the analysis on the basis of the VARMA-model is represented in figure 1.

4. Results and discussions

The elaborated algorithm was tested through the example of modeling price increment for Apple, Microsoft Corporation and Netflix shares.

4.1. Forming a primary statistical information base;
Information on two complete trading sessions from 9.30 to 16.00, September 8 and 9, 2016, for Apple, Microsoft and Netflix with the interval of 300 seconds, was taken as original data for the study. These companies are included in NASDAQ 100 index and they are related to blue chips. Apple is one of the leading high-tech companies, Microsoft is a leader in the field of software engineering and Netflix is a leading entertainment Internet-service in the world.

4.2. Primary analysis of original data
The average price, according to Apple quotations, September 8, 2016, starts at the highest level in comparison with Netflix and Microsoft quotations – 107 US dollars – and flattens out to 106 US dollars during the day. The analogous trend is for Netflix shares – after some fall at the beginning of trading the price flattens out before the end of the trading day.

Dynamics of temporal variations of the average price for quotations of blue chips, Microsoft, during two trading days has smaller spread of values in comparison with the dynamics of quotations of Apple and hovers around 2 dollars.
4.3. Translating statistical data into stationary form
By applying Box-Jenkins method we get the dynamics of price increment for security papers.

Generally, we can note that the greatest dynamics of price increment is observed at the beginning of a trading session. This is characteristic of the shares of three companies being analyzed.

4.4. Identification of a VARMA-model
The most important stage for undertaking a study on the basis of the VARMA-model is identification or selection of a model. At this stage a researcher has to determine which of the models (that is, which of the parameter values, p and q) is suitable for the experiment, on the basis of his interpretation of original values.

A VARMA-model (1, 1) for three variables – Apple, Microsoft Corporation and Netflix shares – in matrix form is presented below (tables 1, 2).

In accordance with these models data one can note strong influence of historical indicators of each company on itself. One can notice that paying attention to the matrix diagonal. Thus, for example, if the price increment of Apple shares increases by 1% in the time period t-1, then the value for shares in the next period will increase by 0.941%.
Table 1. Estimation of parameters of А1 matrix of a VARMA-model (1, 1).

| Security papers       | Apple (t-1) | Microsoft Corporation (t-1) | Netflix (t-1) |
|-----------------------|-------------|----------------------------|---------------|
| Apple (t)             | 0.941       | 0.045                      | -0.025        |
| Microsoft Corporation (t) | -0.002     | 0.981                      | -0.004        |
| Netflix (t)           | -0.012      | 0.012                      | 0.973         |

Table 2. Evaluation value of the VARMA-model constant (1, 1).

| Security papers        | Apple       | Microsoft Corporation     | Netflix     |
|------------------------|-------------|---------------------------|-------------|
| Constant               | 6.133       | 1.646                     | 3.348       |

4.5. Diagnostics of the model – checking the model for robustness and evaluation of the result fidelity

In order to check the model for robustness a method of data grouping was applied – the original time series were divided into two equal time spans – morning and afternoon. The first period takes into account trading data from 9.30 to 12.45, the second one – from 12.46 to 16.00 (tables 3, 4).

Table 3. Estimation of parameters of А1 matrix of a VARMA-model (1, 1), morning.

| Security papers       | Apple (t-1) | Microsoft Corporation (t-1) | Netflix (t-1) |
|-----------------------|-------------|----------------------------|---------------|
| Apple (t)             | 0.888       | 0.066                      | -0.088        |
| Microsoft Corporation (t) | 0.057     | 0.942                      | 0.049         |
| Netflix (t)           | 0.041       | -0.019                     | 0.888         |

Table 4. Estimation of parameters of А1 matrix of a VARMA-model (1, 1), afternoon.

| Security papers       | Apple (t-1) | Microsoft Corporation (t-1) | Netflix (t-1) |
|-----------------------|-------------|----------------------------|---------------|
| Apple (t)             | 0.944       | 0.129                      | -0.056        |
| Microsoft Corporation (t) | 0.002     | 0.867                      | 0.020         |
| Netflix (t)           | -0.023      | 0.017                      | 0.893         |

Applying the model to grouped data has demonstrated that the values of parameters differ, however, their variation is not really strong and the model can be used in the further analysis.

In order to estimate fidelity of the result, that is to estimate the accuracy of developing a modeled result to the real one, a graphic method of applying real dynamics of price increment upon the modeled dynamics, was used (figure 2).
Figure 2. Modeled prices increment compared with the real prices increment for Apple shares on the basis of a VARMA-model (1, 1).

As we can see at the figure, the modeled prices increment doesn't match the real one; there are large differences which indicate that the present VARMA-model modification can't be used for making forecasts for future periods.

5. Conclusions
As there exists strong variation in the values of changes for the real increment and the modeled one, the model has to be improved.

It can be done by selecting other values of p and q parameters of the VARMA-model; also it is possible to use trading data with another frequency – 60 or 600 seconds – for the analysis.

Besides, a VARMA-model of several orders can be used, that is, taking into account the influence of historical data and a computation error possibility of not only one previous period but of several of them.

Scientific novelty of the present study is that an algorithm for developing a VARMA-model is suggested in it (this model is used for analyzing high-frequency time series and is not presented in the national specialized literature) and the suggested algorithm is tested through the example of modeling prices increment for Apple, Microsoft Corporation and Netflix shares, listed on the NASDAQ exchange. The work contributes to financial econometrics and statistics related to the analysis of high-frequency time series.

References
[1] Shevchenko O 2015 Novelties of Russian Legislation and Actual Problems: College Textbook (Moscow: Prospect)
[2] Gratham B 2016. Analysis of Security Papers (Moscow: Wiliams)
[3] Schwager D 2018. Technical Analysis (Moscow: Alpina Publisher)
[4] Campbell H 2011 NASDAQ journal 8 23–37
[5] Granger C 1969 STOR 3 424–438
[6] Tsay R 2002 Analysis of Financial Time Series (New Jersey: John Wiley & Sons, Inc.)
[7] Tiao G 1985 Biometrica 2 299–315
[8] Lütkepohl H 2005 New introduction to multiple time series analysis (Berlin: Springer-Verlag)
[9] Johansen S 1988 Journal of economic dynamics and control 3 231–254
[10] Kennedy P 1998 A Guide to Econometrics (New Jersey: WileyBlackwell)