sciendo

Trading using Hidden Markov Models during COVID-19 turbulences

Iulian Cornel LOLEA
Bucharest University of Economic Studies, Bucharest, Romania
loleaiulian@gmail.com

Simona STAMULE
Technical University of Civil Engineering of Bucharest, Bucharest, Romania

Abstract. *Obtaining higher than market returns is a difficult goal to achieve, especially in times of turbulence such as the COVID-19 crisis, which tested the resilience of many models and algorithms. We used a Hidden Markov Models (HMM) methodology based on monthly data (DAX returns, VSTOXX index Germany's industrial production and Germany's annual inflation rate) to calibrate a trading strategy in order to obtain higher returns than a buy-and-hold strategy for the DAX index., following Talla (2013) and Nguyen and Nguyen (2015). The stock selection was based on 26 stocks from DAX's composition, which had enough data for this study, aiming to select the 15 best performing. The training period was January 2000 - December 2015, and the out-of-sample January 2016 - August 2021, including the period of high turbulence generated by COVID-19. Fitting the best model revealed that the following regimes are the most suitable: two regimes for DAX returns, two regimes for VSTOXX and three regimes for the inflation rate and for the industrial production, while the posterior transition probabilities were event-depending on the training sample. Furthermore, portfolios built using HMM strategy outperformed the DAX index for the out-of-sample period, both in terms of annualized returns and risk-adjusted returns. The results were in line with expectations and what other researchers like Talla (2013), Nguyen and Nguyen (2015) and Varenius (2020) found out. We managed to highlight that a strategy calibrated based on HMM methodology works well even in periods of extreme volatility such as the one generated in 2020 by COVID-19 pandemic.*

Keywords: COVID-19, Markov chains, financial markets, algorithm, trading, volatility.

Please cite the article as follows: Lolea I.C. and Stamule S. (2021), "Trading using Hidden Markov Models during COVID-19 turbulences", *Management & Marketing. Challenges for the Knowledge Society*, Vol. 16, No. 4, pp. 334-351, DOI: 10.2478/mmcks-2021-0020.

Introduction

Stock prices forecasting is a challenging topic, but an intensely debated one in the literature. This problem is complex due to volatility, seasonality and time dependence of the data, especially when referring to a period of extreme turbulence as it was most of 2020, which was marked by economic lockdown and very high volatility in financial markets due to COVID-19 pandemic. The negative effects of the COVID-19 pandemic have been felt in

several sectors: industry (e.g. Carrasco-Sierra et al., 2020); tourism (Machová et al., 2021) and others.

This research can be extremely useful for retail investors in order to build a better-performing portfolio than an ETF that tracks the general stock market index in Germany (DAX). At the same time, we aimed to reach a better risk management within the portfolio, targeting volatility controlling and to consider transaction costs, which could drag down significantly the performance for a retail investor. The proposed strategy takes into account microeconomic and macroeconomic variables in determining the companies to be part of the portfolio, offering a broad perspective on the factors that can influence stock market performance. Thus, the analysis could also be useful for institutional investors with more complex investment strategies that can incorporate it into their model, taking into account the buy or sell signals provided by this HMM strategy.

There are classic forecasting models such as the autoregressive ones that can work with a certain level of accuracy. More recently, a series of artificial intelligence algorithms was developed, such as neural networks, random forest or SVM (support-vector machine), as Tudor (2014) applied in order to make predictions for the Romanian equity market. However, there are a lot of limitations of these methodologies. The risk of overfitting is present, while some of them have the major disadvantage of being so-called "black-boxes" without knowing exactly how they work and without being able to introduce the knowledge and flair of market traders. Given the disadvantages of these and the fact that it has been proven that models based on hidden Markov Models can be used successfully for modeling time-dependent phenomena and for making predictions on the equity market, we aimed to analyze the applicability of such models and to highlight their usefulness.

The purpose of this paper is to calibrate the HMM-based model using as variables the DAX returns, the German inflation rate and the VSTOXX index as a proxy for stock market volatility and the monthly change of industrial production. We chose VSTOXX index because we could not identify a closer proxy for equity market implied volatility in Germany. Based on these data, the model will be calibrated for a period of 16 years. We will perform tests and analyzes and subsequently, we will test its' applicability out-of-sample (January 2016 - August 2021) to verify whether it brings added-value versus a buy-and-hold approach. The strategy will also consider trading costs to be as close as possible to reality. This approach is going to be different from that of Nguyen and Nguyen (2015). We expect the results to lead to better performance than a buy-and-hold strategy, as other authors have achieved, even if testing is done for a period of high turbulence generated by COVID-19 (Landmesser, 2021; Waliszewski & Warchlewska, 2021). We also expect the estimation of posterior transition probabilities for training sample to indicate that the model is properly calibrated to cope with periods of high volatility, given that important events such as the dot.com crisis, the economic and financial crisis of 2008 or the sovereign debt crisis are included.

The period for which the performance of the model will be tested is January 2016 - August 2021, this being a period that includes the major turbulences generated by the lockdown due to COVID-19 pandemic, and by the problems generated in the economy (e.g. Gavurova et al., 2021). We aimed to include in the out-of-sample period years in which the evolution of the equity market was not affected by major events, but also years in which there were strong turbulences to test the viability of the model. At the same time, the economy during the lockdown period was confronted with unprecedented decisions that

had an impact on the financial markets and brought important changes in investor behavior. The literature is still poor in terms of research on the impact of COVID-19 on financial markets (Pardal et al. 2020; Korzeb & Niedziółka, 2020; Liu, N. et al., 2021), and among these, we consider worth mentioning Cepoi (2020) which analyzes the dependence between stock markets and new information on COVID-19.

The application of this strategy for a period of high turbulence in the stock markets (2020-2021) following the COVID-19 pandemic confirmed the viability of the model and its applicability. At the same time, the usefulness in the efficient portfolio management was revealed by obtaining better results, both from the point of the annualized return and from the point of risk-adjusted returns. Thus, the crises generated by COVID-19 and the lockdown period were a test for the model in periods of high turbulence.

The paper also has an important contribution to the scientific literature. It evaluates the performance of a strategy based on HMM models during COVID-19 crisis period in 2020 and discusses the use of a variety of risk indicators to make comparisons. Also, unlike Varenius (2020) and Talla (2013), who conducted studies for the European market, this paper uses the VSTOXX index for implicit volatility, proving its usefulness, as in the case of VIX. Moreover, the estimation of unconditional probabilities provides a visual perspective on the regimes considered in the model.

The rest of the paper is organized as follows. Section 1 is the presentation of the existing literature related to our topic. Section 2 is dedicated to methodology and data used in the paper and defines what methods we have used, introducing the reader to the Markov chains and Markov models. Section 3 focuses on choosing most fitted model and testing the performance of this model against a buy and hold strategy. The next section is dedicated to conclusions and at the end of this paper you can find the references of this research.

Literature review

HMM models can be applied directly for equity market data, but can be used even more complexly by integrating more information while making a connection between the equity market and macroeconomic variables that can influence the dynamics of financial markets. This approach was taken by Talla (2013) for the Swedish equity market, using the OMXS30 index with encouraging results. Moreover, we have considered another three important researches for the applicability of HMM in order to forecast equity markets using macroeconomic variables: Nguyen and Nguyen (2015), Ryou et al. (2020), Varenius (2020), and by Nguyen and Nguyen (2021).

Based on these ideas and previous research, our purpose in this article was to use the application of HMM to highlight whether it is possible to build a model that achieves significantly better trading results than a buy-and-hold strategy. This analysis will be applied to the DAX index and for companies that are part of it. This is a total return index, for which more relevant conclusions can be drawn over time in terms of performance. Papers that have applied similar methodologies to the German equity market are those of Hassan et al. (2013) or Oelschläger and Adam (2020).

The basis for modelling data using HMM is represented by papers focused on classical Markov models. Of these, the work of Marcucci (2005) is representative in that he applied Markov Switching to identify changes in the volatility regime of the United States market. He used the volatility regime switching to improve the performance of different types of GARCH models. Marcucci's (2005) approach was subsequently improved and

extrapolated to different markets, as Fei et al. (2017) did by combining the Markov chain methodology with the Copula function to model the dependency between the CDS (credit default swap) market and the equity market. Also, De la Torre et al. (2018) and He et al. (2018) extended the Markov Switching approach to different asset classes such as commodities. Moreover, they developed comparisons regarding the usefulness of Markov chains for regime switching in emerging markets vs. developed markets.

In order to substantiate the working methodology within this analysis, we went through a series of existing works, which we considered useful for our approach. We noticed that HMM (Hidden Markov models) was applied at the beginning for solving problems other than financial ones such as the recognition of different patterns, handwriting or gestures, as applied by Vieriu et al. (2011). Subsequently, this methodology was extrapolated to financial markets to recognize different patterns and to be used in making forecasts. One of the first papers in this regard was that of authors Rydén et al. (1998), followed later by Hassan and Nath (2005). The latter made a comparison in terms of performance of HMM-based models with those based on artificial neural networks, and the results were nearly the same, proving that there is significant progress in implementing HMM for financial markets.

We follow Nguyen and Nguyen (2015), Kim et al. (2019), Varenius (2020) and Nguyen and Nguyen (2021) who aimed to forecast the evolution of returns on equity markets using both historical indices data and macroeconomic variables, further developing the approach of Talla (2013), which considered that it is possible to achieve an effective forecast by introducing significant macroeconomic variables in the models. Among these variables, they also had into attention the inflation rate, the yield curve of government securities or corporate bonds, the money market interest rates, the CDS' curve or the commodities' prices. These are variables that have a higher frequency and can be used in forecasts of stock market movement. For each of these articles, the results were gratifying and proved that models based on hidden Markov chains can add value and can be the basis for useful forecasts.

Other important papers used in documenting the methodology are the ones published by Dima and Vasilache (2020), Ryou et al. (2020), Varenius (2020) and Chandrika et al. (2020) which applied exactly these methods based on HMM and financial market variables, but also on macroeconomic variables in order to obtain equity market forecasts. Varenius (2020) applied the methodology for the OMXS30 index and achieved better results than a buy-and-hold strategy, considering factors that could distort the results such as: corporate actions, dividend yield, trading and portfolio management costs. The same type of approach was previously taken by Cuellar Andersson and Fransson (2016), also for the Swedish equity market, aiming to obtain a trading model that has better results than the OMXS30 index. The methodology presented by these authors is a well-known one in the literature and will be the basis of our study.

The applicability and usefulness of HMM-based methodologies has been proven by comparisons between machine learning and HMM-based methods. This approach was applied by Nabipour et al. (2020) focusing on deep-learning models or by Liu et al. (2021), focusing on the LSTM methodology. Furthermore, we identified two reference studies (Uysal and Mulvey (2021), Zhang et al. (2021)) that improve the machine learning approach for stock price prediction using Markov chains and 2 volatility regimes. This improvement in methodology has proven to add value to using machine learning for a

single volatility regime. Thus, the utility of Markov prose and HMM models was proven once again, even in combination with other well-known methodologies.

An approach of great interest, but which our study touches only very little for now, is that of exploiting the adaptive market hypothesis (AMH), as Burhan and Eylem (2021) did in their paper, where they used HMM models to be able to interpret the evolution of these markets. In the same sense, we can mention the work of Waduge and Ganegoda (2018), who followed the idea of adaptive markets (AMH) and built models that include macroeconomic variables with higher frequency in order to make certain forecasts on equity markets, trying to benefit from periods where markets were inefficient.

Materials and methods

Markov models

The methodology regarding Markov chains and Hidden Markov Model is a very complex one and we have followed an online published course on Markov Models by Jurafsky and Martin (2021) for Stanford University. This methodology was also applied by Nguyen and Nguyen (2015) and Varenius (2020). Moreover, I have used the paper of Visser and Speekenbrink (2010) which describes very well the Hidden Markov methodology and how to use depmixS4 package in RStudio to implement it.

A model with hidden Markov chains is represented by two connected random processes: a hidden process $\{X_k\}$ and an observable process $\{Y_k\}$, which are connected in the idea that $\{X_k\}$, determines the distribution of the corresponding variable $\{Y_k\}$. In this sense, it can be stated that if $\{X_k\}$, follows a normal distribution, it influences the mean and the variance of $\{Y_k\}$, so that it also has a normal distribution.

This connection was explained in the following way: for the current state the hidden process is independent of the observed process. The observed process depends on the past values and the hidden process.

As Jurafsky and Martin (2021) stated in their course, a Markov Hidden Model can be characterized by five elements:

- There is a finite number of states in the model;
- There is a finite number M of symbol observations for each state. The individual symbols will be denoted V = {v_1, v_2… .v_M};
- There is a transition matrix of probabilities between states denoted A = {a_ (i, j)}, where the values inside it are defined as follows:

$$a_{ij} = P\left(q_{t+1} = S_j | q_t = S_i\right), \;\; 1 \le i, j \le N \qquad (1)$$

- There is a probability distribution in state j, $B = \left\{b_j(k)\right\}$, where the elements in the matrix are defined as follows:

$$b_j(k) = P\left(v_k \; la \; t | q_t = S_j\right), \;\; 1 \le j \le N, \;\; 1 \le k \le M. \qquad (2)$$

- There is an initial distribution of states $\pi = \{\pi_i\}$,

Where $\pi_i = P(q_1 = S_i), \;\; 1 \le i \le N$.

Therefore, a complete description of the HMM model requires the following: a specification of the two model parameters (N and M), specification of the observation symbols and the three probability measures A, B and π. Next, I will use the following notation for simplification:

$$\tau = (A, B, \pi)$$

The sequence of observations will be denoted by $O = \{O_1, O_2 \ldots . O_T\}$.

According to Jurafsky and Martin (2021) there are three important issues that can be defined when estimating an HMM model:

a) Likelihood: how the probability that the sequence of observations O is produced by the parameters in τ should be properly calculated. This probability will be denoted by $P(O|\tau)$.
b) Decoding: this involves the optimal choice of the sequence of states $Q = \{q_1, q_2, \dots., q_t\}$ that best describes the observations to which it applies.
c) Learning: how to adjust the parameters A, B, π in order to maximize $P(O|\tau)$.

To solve the problem from point "a)" the Backward or Forward algorithms will be applied and a result will be reached regarding the parameters to be estimated.

To solve the problem in point "b)", one can apply the Viterbi algorithm. The solution for this algorithm is reached by an iterative process.

To solve the problem in point "c)", according to the theory introduced by Rabiner (1989), we will apply the Baum-Welch algorithm. This algorithm assumes that the parameters A, B, τ will be adjusted to maximize P (O | τ). According to the theory, this is the most difficult part of estimating an HMM model, being known that there is not an analytical method to reach a solution. The Baum-Welch algorithm uses an iterative procedure that chooses τ = (A, B, π) to find a local maximum for P (O | τ). This approach is based on the Expectation-Maximization algorithm.

After estimating the most appropriate number of regimes for each data series, we will proceed to the stock selection for each month. We will check for the training period, which were the performing 15 stocks each scenario and we will select them for the following month. Furthermore, we will not use equally-weighted portfolio as other authors like Varenius (2020) and Nguyen and Nguyen (2021) did. We will implement the following methodology to determine the weight for each stock in order to better control portfolio's risk:

- Ranking the stocks by their monthly volatility for the last year (the higher de volatility the lower the grade);
- Each stock got a grade from 1 to 15;
- Based on these grades, we have the new weights for each stock in portfolio according to the following formula:

$$weight\ of\ stock\ i = \frac{grade\ of\ stock\ i}{\sum_{j=1}^{15} grade\ of\ stock\ j} \quad (3)$$

Portfolio's performance will be measured by the following indicators:

- Annualized return;
- Annualized volatility;
- Upside standard deviation;
- Downside standard deviation;
- VaR (95%);
- Maximum Drawdown.

The data

The study is based on 4 data series: monthly return of DAX, annual inflation rate in Germany industrial production in Germany, VSTOXX index which reveals the implied volatility for Euro Stoxx 50. For all four data series the frequency was monthly, this being

the highest frequency for which there are common data. These data were extracted from public data sources for the period between January 2000 - August 2021.

Alongside to these data, we have also used information regarding 26 companies from the structure of DAX for which data could be identified for the entire analyzed period Jan. 2000 – Aug. 2021. These companies were: Adidas, Airbus SE, Allianz SE, BASF SE, Bayern AG , BMW AG, Continental AG, Daimler AG, Dutsche Bank AG, Deutsche Borse AG, Deutsche Telekom AG, E.ON SE, Fresenius Medical Care AG, HeidelbergCement AG, Henkel AG, Infineon Technologies AG, Merck KGaA, Muenchener Rueckversicherungs-Gesellschaft AG , Porsche Automobile, Puma SE, QIAGEN NV, RWE AG, SAP SE, Siemens AG, and Volkswagen AG.

It should be noted that DAX index is a total return index (i.e. it includes dividend yields) and, in this way, for comparability, we obtained dividend-adjusted prices for the 26 companies. Further, we will present descriptive statistics for the main 3 data series used for calibrating the model.

Table 1. Descriptive statistics

Indicator	VSTOXX	DAX Returns	Inflation	Industrial production
Mean	23.4	0.5%	1.45%	0.06%
Median	21.3	0.8%	1.40%	0.20%
Minimum	13.2	-25.4%	-0.5%	-21.33%
Maximum	79.1	21.4%	4.1%	11.14%
Skewness	1.73	-0.55	0.17	-2.72
Excess of Kurtosis	4.15	2.23	0.96	32.27
Observations	262	262	262	262

Source: authors' own computations.

These data will be divided into two sub-samples: Training data: Jan 2020 - dec 2015 and testing data (out-of-sample): Jan. 2016 - Aug. 2021.

Results

It is worth mentioning that the same approach was conducted by Talla (2013), Nguyen and Nguyen (2015) and by Varenius (2020). However, they used different variables, did not consider trading costs, and did not consider unconditional posterior probabilities, which we find very important to follow. Also, our model will test the viability of HMM during COVID-19 Pandemic, which was a period of high turbulences and will introduce a scheme of weighting the stocks inversely proportional to the volatility. Furthermore, we think the metrics for performance measurement in the previous papers were not enough, hence we introduced new indicators such as downside standard deviation, upside standard deviation or VaR (95%). These new metrics may bring some new information regarding portfolio performance in different periods of time.

When estimating the most relevant number of states we have stopped at a maximum of 4 states (regimes) to avoid overfitting and because there is an economic justification only for this maximum number of states. Thus, for the equity market, there is an intuitive justification for two states (high volatility or low volatility). On the other hand, for macroeconomic variables there may be intuitive justifications for several conditions such

as: periods of inflation, deflation or disinflation. We can also talk about galloping inflation or hyperinflation. On the other hand, economic cycles can be divided into four phases: growth, boom, recession and depression. Therefore, we can say that in terms of the number of regimes, there could be a justification up to a maximum of four states, according to economic theory and intuition, and these will be tested and verified later.

In the first part of this chapter we will present the estimations' results and how we arrived at the most appropriate model in order to achieve a better performance than a buy-and-hold strategy. The first data series for which the HMM methodology was applied was represented by the monthly DAX returns for the training period: Jan. 2000 – Dec. 2015. Estimating the results for the DAX returns using 2, 3 and 4 regimes led to the following values for AIC and BIC.

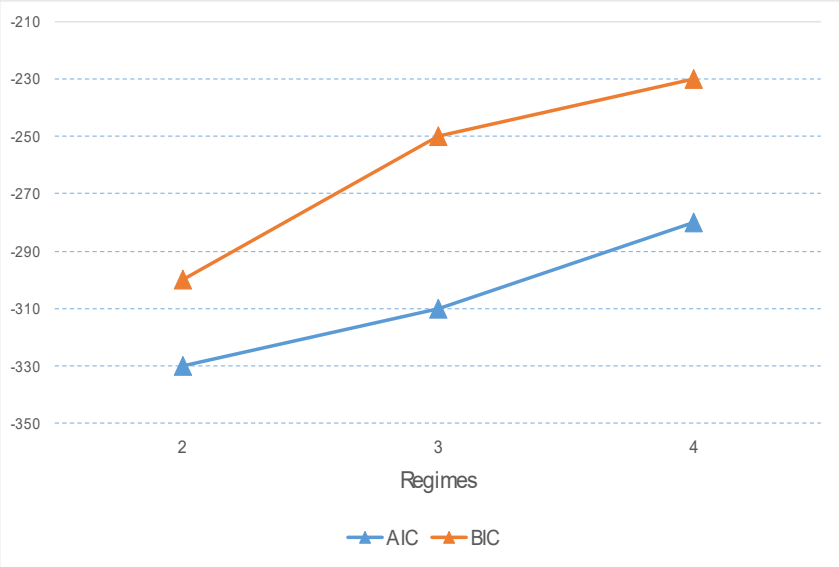


Figure 1. AIC and BIC values for the HMM model based on DAX returns

Source: authors' own computations.

The information criteria indicate that the most suitable model for DAX returns is the one with 2 regimes. Thus, in the trading algorithm we will use the 2-regime model for DAX index returns. This result is in line with what Marcucci (2005), Talla (2013), Nguyen and Nguyen (2015) and Varenius (2020) found out.

We will further present the detailed results for the 2-states model:

Table 2. 2-states model for DAX returns

Regime	Observations	Mean	Volatility
1	**79**	**-1.16%**	**6.3%**
2	**112**	**1.57%**	**5.1%**

Source: authors' own computations.

The results presented in the previous table show that the regime number 1 is the one related to higher volatility, i.e. the one related to the periods of decline on the equity markets, having an average monthly return of -1.16%. On the other hand, the second regime is the one for lower volatility and the one for the periods of positive returns with a

monthly average of 1.57%. It should also be noted that the regime with lower volatility has 112 observations compared to 79 for the high-volatility regime.

In order to reveal the regime changes, we have presented in the following graph the evolution of posterior transition probabilities, which were not taken into consideration in most of former papers on this topic:

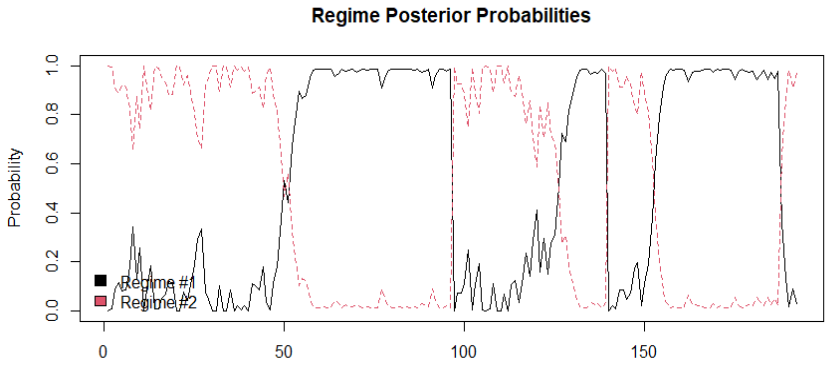


Figure 2. Unconditional posterior probabilities for the DAX model with 2 regimes

Source: authors' own computations.

The second variable for which we have estimated the models with 2, 3 and 4 states is the VSTOXX index. The values obtained for AIC and BIC are presented in the following graph:

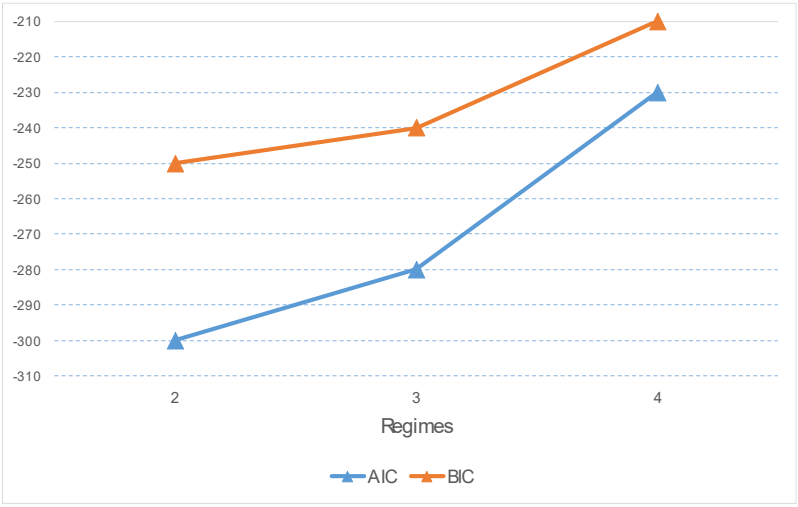


Figure 3. AIC and BIC values for the HMM model based on VSTOXX Index

Source: authors' own computations.

The previous figure showed that the most suitable model is the one with two states, as for the DAX returns, taking into consideration that AIC and BIC are minimum criteria.

The detailed results for the case of two-state HMM model are presented below:

Table 3. 2-states model for VSTOXX Index

Regime	Observations	Mean	Volatility
1	93	16.4	2.8
2	98	25.6	5.4

Source: Authors' own computations.

The first regime is the one with a lower volatility, for a number of 93 observations, and the second one is with a higher volatility, for a number of 98 observations, with an average value of the VSTOXX index of 25.6. For these results we also presented the dynamics of unconditional transition probabilities:

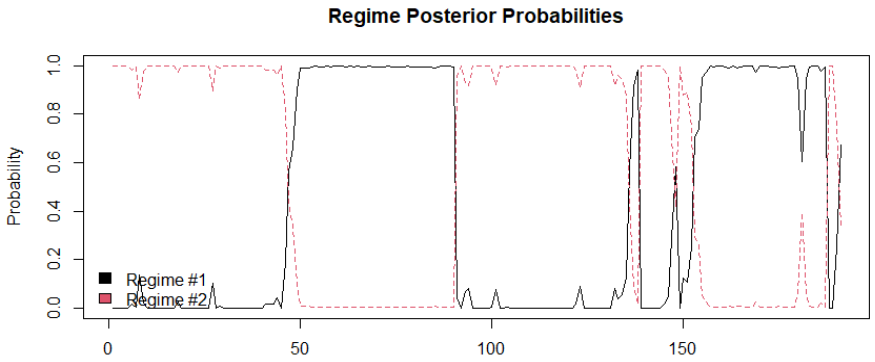


Figure 4. Unconditional posterior probabilities for the VSTOXX model with 2 regimes

Source: authors' own computations.

The third variable for which the estimate was made is the annual monthly inflation rate for Germany. We proceeded in the same way and calculated the values for the information criteria AIC and BIC, and the following graph shows the results obtained:

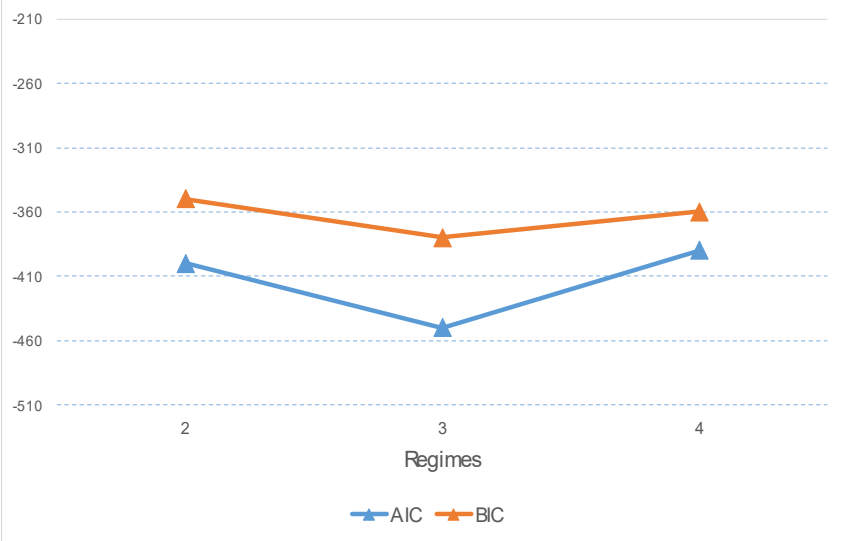


Figure 5. AIC and BIC values for the HMM model based on the inflation rate

Source: authors' own computations.

The values for the two informational criteria indicate that the most appropriate model for this data is the one in which 3 regimes were used. As in the case of the other variables, we will present in detail only the results for the chosen model:

Table 4. 3-states model for the inflation rate

Regime	Observations	Mean	Std. Dev
1	93	1%	0.9%
2	43	2.20%	0.5%
3	55	0.80%	0.3%

Source: authors' own computations.

It can be seen that there is a regime with average inflation, an intermediate regime (regime 3). It has an average inflation rate of 0.8%. The results indicated an increased persistence for all three regimes.

For a better illustration of the results we will also present the evolution of the unconditional probabilities for the three regimes related to the inflation rate, according to the HMM model used and estimated in R-Studio.

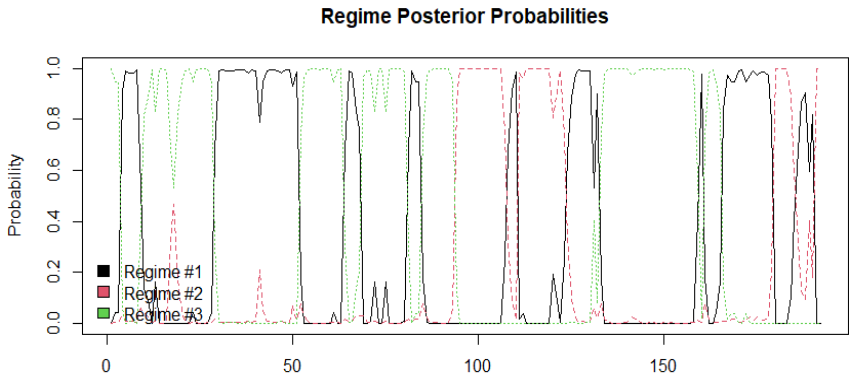


Figure 6. Unconditional posterior probabilities for the inflation rate model with 3 regimes

Source: authors' own computations.

The last variable for which the estimate was made is the monthly change of industrial production in Germany. We proceeded in the same way and calculated the values for the information criteria AIC and BIC, and the following graph shows the results obtained:

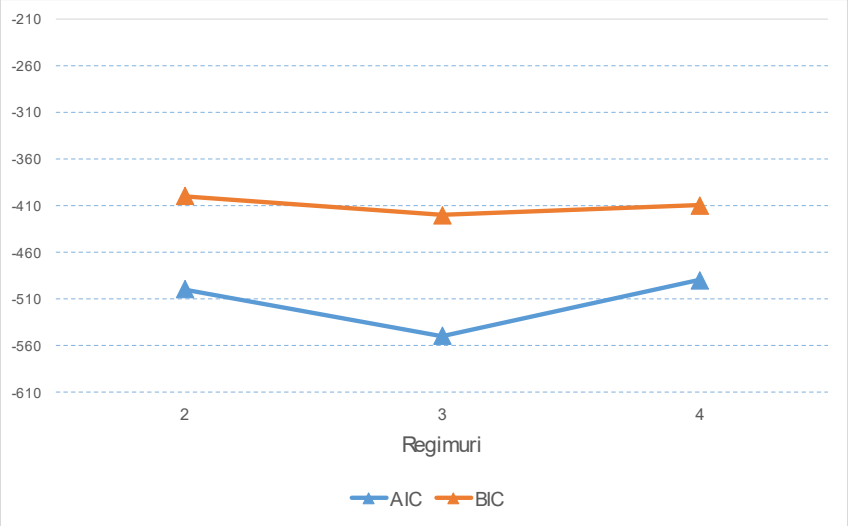


Figure 7. AIC and BIC values for the HMM model based on the industrial production

Source: authors' own computations.

As in the case of inflation rate, the AIC and BIC criteria indicated that the most appropriate model is the one with 3 regimes. Further, we will present detailed results for the 3-states estimations:

Table 5. 3-states model for the inflation rate

Regime	Observations	Mean	Std. Dev
1	93	0.40%	2.4%
2	55	2.30%	4.5%
3	43	-1.40%	6.5%

Source: authors' own computations

Regime 1 generated most observations (93) and this seems to be the regime for growth period with slow and moderate growth. On the other hand, the second regime seems to be the regime for the boom period and the third one seems to be the regime for the period of recession.

In order to complete the presentation, we will also illustrate de evolution of posterior unconditional probabilities for industrial production:

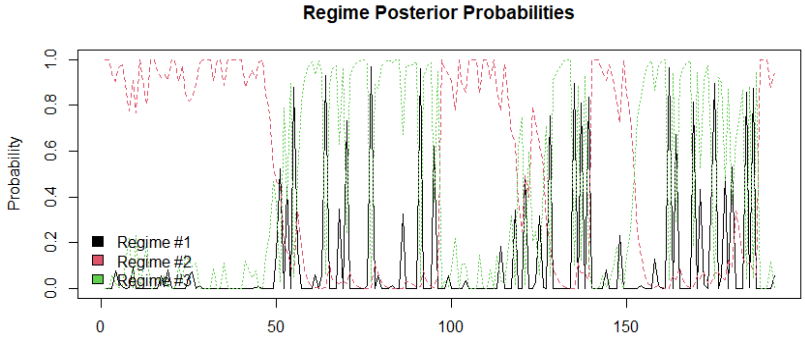


Figure 7. Unconditional posterior probabilities for the industrial production with 3 regimes

Source: authors' own computations.

Based on the results obtained for the training period of the model (Jan. 2000 - Dec. 2015), we went further and made estimations for the interval used as out-of-sample (Jan. 2016 - Aug. 2021). The number of regimes used comes in the following order: (2, 2, 3, 3) - DAX, VSTOXX, inflation rate and the monthly change in industrial production.

For each month in which we forecast a certain regime for each variable, we checked during the training period which were the first 15 stocks that performed best in such a scenario and we selected them for the following month. Using this approach for each month, we managed to get a trading methodology. Furthermore, we did not use an equally-weight portfolio as Nguyen and Nguyen (2015) and Varenius (2020) did. We have determined a weight for each stock based on the above methodology, in order to better control for taking too much risk:

- We have ranked the stocks by their monthly volatility for the last year (the higher de volatility the lower the grade);
- Each stock has a grade from 1 to 15;
- Based on these grades we have calculated the weight for each portfolio:

$$weight\ of\ stock\ i = \frac{grade\ of\ stock\ i}{\sum_{j=1}^{15} grade\ of\ stock\ j} \quad (4)$$

This may be a methodology that could be properly used for constructing low volatility indices.

Following all this methodology, we have built two different portfolios to cover two different approaches that we have identified in the literature:

- Portfolio in which trading costs were not taken into account, according to the approach taken by Nguyen and Nguyen (2015) and Nguyen and Nguyen (2021).
- Portfolio in which trading, selling and buying costs were taken into account, when changes were made to certain stocks in the portfolio compared to the previous month, according to the approach taken by Varenius (2020) and Ryou et al. (2020).

We specify the fact that, for the transaction costs, the values that a Romanian retail investor could receive from a trading broker were taken into account. We considered an average trading fee of 0.35% per transaction. If this approach was applied by an institutional investor, with higher trading volumes and access to preferential commissions, the trading cost could fall even up to 0.1% per transaction, making the model more profitable.

The evolution of the two portfolios for the training period is presented below. We also highlighted the way in which the two portfolios and DAX behaved during the crisis generated by the COVID-19 pandemic, this being the interval for which the maximum drawdown was registered.

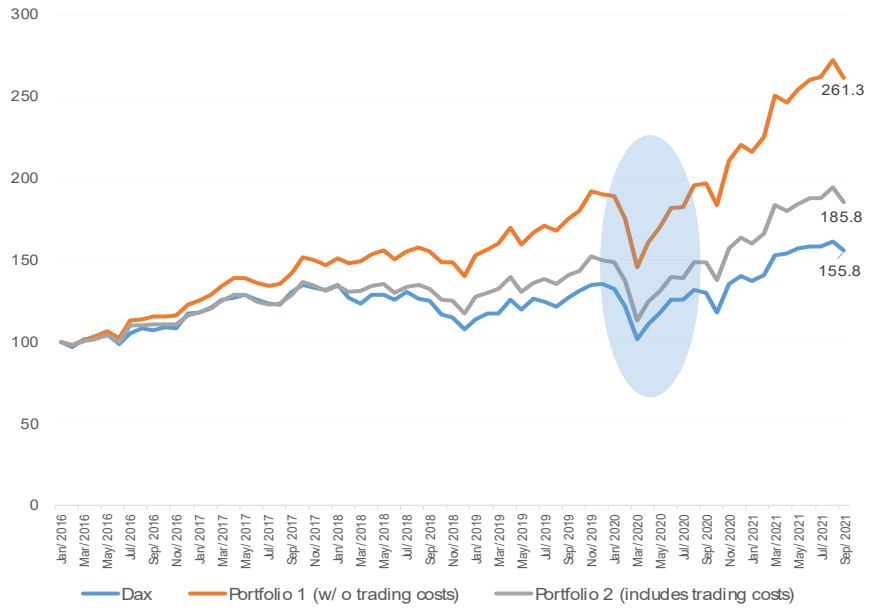


Figure 8. Portfolio 1, Portfolio 2 vs. DAX – performance evolution

Note: The 2 portfolio and DAX were normalized at 100 for comparability.

Source: authors' own computations.

The previous figure shows DAX's evolution and the two portfolios between Jan. 2016 and Aug. 2021. The results indicated a much better performance of the portfolios built according to the rule described in this paper. It is also noted that the portfolio that does not include trading costs performed 161.3%, significantly above DAX (55.8%) and above the portfolio which takes into consideration transaction costs (85.8%).

The centralized results for the three portfolios are presented below, using a table:

Table 5. Risk and return performance indicators

Indicator	**Buy-and-Hold**	**Portfolio 1[1]**	**Portfolio 2[2]**
Return (%, p.a.)	8.1%	18.4%	11.5%
Annualized volatility	16.9%	17.9%	17.8%
Upside Std. Dev.	12.4%	13.3%	13.2%
Downside Std. Dev	20.5%	21.4%	21.5%
Sharpe ratio	0.42	0.97	0.59
VaR (95%) -	27.8%	29.4%	29.3%
Maximum drawdown	-25.1%	-17.2%	-17.6%

[1]This portfolio does not include trading costs; [2]This portfolio does include trading costs.

Source: authors' own computations.

Based on these results, it can be seen that the HMM-based strategy adds value to most of the indicators considered for analysis. Firstly, in terms of annualized return, we registered an increase of 2.4 p.p. compared to the buy and hold strategy.

We used maximum drawdown and annualized volatility to measure the risk. For the first indicator, the HMM-based strategy performed better, while, for annualized volatility the two portfolios registered slightly higher values than a buy-and-hold strategy. The same observation can be made for the annual VaR (95% - Value at Risk).

Moreover, we considered useful to calculate a risk-adjusted return indicator (Share Ratio), given that the two types of strategies performed differently in terms of risk. In the case of the buy-and-hold strategy we recorded a value of 0.42, while for the portfolio built according to HMM and which includes trading costs, we obtained a value of 0.59. These values suggest that the implementation of HMM methodology added value compared to a classical approach of passive portfolio management.

Conclusion

Through this study, we aimed to apply the HMM methodology to obtain a trading strategy for companies that are part of the DAX index. Our study used monthly data for DAX, VSTOXX, Germany's annual inflation rate, Germany's monthly change of industrial production and 27 companies, which are constituents of the DAX index. The data set consists of two sub-periods: i) Jan 2000 – Dec. 2015 – training period, ii) Jan 2016 – Aug 2021 – testing period – out-of-sample, which included the crisis period generated by COVID-19 pandemic.

The implementation of HMM methodology led to the following results regarding the fitting of the most suitable number of regimes: 2 states for DAX returns, 2 states for VSTOXX Index and 3 states for the inflation rate and the monthly change of industrial production. This result is in line with the economic theory and the posterior unconditional probabilities for high volatility regimes correctly indicated the turbulent periods in the training set. At the same time, the result considered that the regimes in the model are in line with Nguyen and Nguyen (2015) and Varenius (2020), which use the same methodology. The number of regimes chosen for the macroeconomic variable is in line with what other authors such as Talla (2013) and Varenius (2020) have achieved. At the same time, the capture of changes in the equity market through 2 regimes is consistent with the results obtained by Nguyen and Nguyen (2015), who applied a similar approach. However, questions arise regarding the accuracy of reflecting all equity market conditions when using low frequency data (monthly), with the risk that certain turbulences will occur for a higher frequency (weekly, daily or even intraday) and we will not capture them.

The results for testing sample indicated an over-performance compared to the buy and hold strategy, both when taking into account the trading costs, as well as when we don't. We consider that the approach in which trading costs are taken into account is more realistic given that the methodology proposed in this paper involves a frequent (monthly) trading of at least 6 shares in the portfolio (the minimum number of shares that have been changed in the portfolio compared to previous month). Thus, we consider that a different approach from that of the authors Nguyen and Nguyen (2015), an approach that considers transaction cost, is a more appropriate one. In this sense, as mentioned before, we considered the transaction costs that are available for a retail investor in Romania. The results were consistent with those recorded by other authors who applied similar methodologies. Among the papers to which we can compare our results, we can mention the articles of Nguyen and Nguyen (2015), Ryou et al. (2020), Varenius (2020) and Nguyen and Nguyen (2021).

Furthermore, it is very important that the strategy has led to better performance in terms of risk, but also in terms of excess return over the risk-free rate or risk-adjusted return (Sharpe Ratio). It is noteworthy that the strategy performed well even if it was a period affected by the COVID-19 crisis and marked by economic lockdown. The results were in line with what other authors such as Talla (2013), Nguyen and Nguyen (2015), Varenius (2020) obtained, even if the period for which we applied the new methodology was one marked by extreme turbulence, due to COVID-19, unlike those in their studies.

On the other hand, the fact that the strategy performed better in the period 2016 - 2021, in which we had periods of low volatility alternating with periods of very high volatility, may suggest that there are chances of obtaining higher returns than the market. Hence, there could be some evidence that a well-known theory like the Adaptive Market Hypothesis could be validated Burhan and Eylem (2021). At the same time, exploiting this theory can be a way to expand the study in the future, including even other methodologies through which one could benefit of market opportunities. Among the methods discussed lately in the literature, those that apply machine learning techniques would be of interest.

The practical usefulness of this case study and how its applicability can be extended should not be neglected. A retail investor can be guided in trading by the signals that the HMM-based strategy can provide and could obtain higher returns than the market, even when considering the trading costs. The return obtained could be improved if the strategy is implemented by an institutional investor who has access to lower trading fees. At the same time, the results of the study can be used in a more complex strategy that includes several types of models or that aims to build early-warning indicators for turbulent periods in the financial markets.

However, the methodology presented has its limitations. Firstly, we are talking about monthly trading and in this case the risk to lose some of the opportunities generated by the movements inside the month arises, even though trading more frequently would imply higher trading costs. Moreover, it uses only a limited number of criteria to compare the strategies. There may be some interest in considering investors' risk tolerance, turning the problem into a behavioral one.

References

Burhan, H. A., & Eylem, A. (2021). Adaptive Market Hypothesis and Return Predictability: A Hidden Markov Model Application in Borsa Istanbul. *Sosyoekonomi*.

Carrasco Sierra, A., Cobos Flores, M.J., Fuentes Duarte , B., Hernández Comi, B.I. (2020). Successful Management System by a Metalworking Mexican Company During COVID-19 Situation. Analysis Through a New Index (Case Study). *International Journal of Entrepreneurial Knowledge*, 8(2), 42-55.

Cepoi, C. O. (2020). Asymmetric dependence between stock market returns and news during COVID-19 financial turmoil. *Finance Research Letters*, 36, 101658.

Chandrika, P. V., Visalakshmi, K., & Srinivasan, K. S. (2020, March). Application of Hidden Markov Models in Stock Trading. *In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 1144-1147). IEEE.

Cuellar Andersson, J., & Fransson, L. (2016). Algorithmic Trading Based on Hidden Markov Models–Hidden Markov Models as a Forecasting Tool When Trying to Beat the Market.

De la Torre, O. V., Galeana-Figueroa, E., & Alvarez-García, J. (2018). Using Markov-Switching models in Italian, British, US and Mexican equity portfolios: a performance test. *Electronic Journal of Applied Statistical Analysis*, 11(2), 489-505.

Dima, A. & Vasilache, S. (2009). ANN Model for Corporate Credit Risk Assessment. Proceedings - 2009 International Conference on Information and Financial Engineering, ICIFE 2009. 94 - 98. 10.1109/ICIFE.2009.33.

Fei, F., Fuertes, A. M., & Kalotychou, E. (2017). Dependence in credit default swap and equity markets: Dynamic copula with Markov-switching. *International Journal of Forecasting*, 33(3), 662-678.

Gavurova, B., Ivankova, V., Rigelsky, M., Přívarová, M. (2020). Relations Between Tourism Spending and Global Competitiveness – an Empirical Study in Developed OECD Countries. *Journal of Tourism and Services*, 21(11), 38-54.

Hassan, M. R., Ramamohanarao, K., Kamruzzaman, J., Rahman, M., & Hossain, M. M. (2013). A HMM-based adaptive fuzzy inference system for stock market forecasting. *Neurocomputing*, 104, 10-25.

Hassan, M. R., & Nath, B. (2005, September). Stock market forecasting using hidden Markov model: a new approach. *In 5th International Conference on Intelligent Systems Design and Applications (ISDA'05)*,192-196.

He, Z., O'Connor, F., & Thijssen, J. (2018). Is gold a Sometime Safe Haven or an Always Hedge for equity investors? A Markov-Switching CAPM approach for US and UK stock indices. *International Review of Financial Analysis*, 60, 30-37.

Jurafsky, D. & Martin, J.H. (2021), Hidden Markvo Model, Stanford Online Course available at: http://web.stanford.edu/~jurafsky/slp3/A.pdf.

Kim, E. C., Jeong, H. W., & Lee, N. Y. (2019). Global Asset Allocation Strategy Using a Hidden Markov Model. *Journal of Risk and Financial Management*, 12(4), 168.

Korzeb, Z., & Niedziółka, P. (2020). Resistance of commercial banks to the crisis caused by the COVID-19 pandemic: the case of Poland. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *15*(2), 205-234. https://doi.org/10.24136/eq.2020.010.

Liu, N., Xu, Z., & Skare, M. (2021). The research on COVID-19 and economy from 2019 to 2020: analysis from the perspective of bibliometrics. *Oeconomia Copernicana*, *12*(2), 217–268. https://doi.org/10.24136/oc.2021.009.

Liu, M., Huo, J., Wu, Y., & Wu, J. (2021). Stock Market Trend Analysis Using Hidden Markov Model and Long Short Term Memory. *arXiv preprint arXiv*:2104.09700.

Machová, R., Korcsmáros, E., Esseová, M., Marča R. (2021). Changing Trends of Shopping Habits and Tourism During the Second Wave of COVID-19 – International Comparison. *Journal of Tourism and Services*, 22(12), 131-149.

Landmesser, J. (2021). The use of the dynamic time warping (DTW) method to describe the COVID-19 dynamics in Poland. *Oeconomia Copernicana*, *12*(3), 539-556. https://doi.org/10.24136/oc.2021.018.

Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics,* 9(4).

Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. IEEE Access, 8, 150199-150212.

Nguyen, N., & Nguyen, D. (2015). Hidden markov model for stock selection. *Risks*, 3(4), 455-473.

Nguyen, N., & Nguyen, D. (2021). Global stock selection with hidden Markov model. *Risks*, 9(1), 9.

Oelschläger, L., & Adam, T. (2020). Detecting bearish and bullish markets in financial time series using hierarchical hidden Markov models. *arXiv preprint arXiv*:2007.14874.

Pardal, P., Dias, R., Šuleř, P., Teixeira, N., & Krulický, T. (2020). Integration in Central European capital markets in the context of the global COVID-19 pandemic. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *15*(4), 627-650. https://doi.org/10.24136/eq.2020.027.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286.

Ryou, H., Bae, H. H., Lee, H. S. & Oh, K. J. (2020). Momentum investment strategy using a hidden Markov model. *Sustainability*, 12(17), 7031.

Rydén, T., Teräsvirta, T., & Åsbrink, S. (1998). Stylized facts of daily return series and the hidden Markov model. *Journal of applied econometrics*, 13(3), 217-244.

Talla, J. T. (2013). Impact of macroeconomic variables on the stock market prices of the Stockholm stock exchange (OMXS30). Jonkoping International Business School, 01-48.

Tudor, N. L. (2014, May). Intelligent system for time series prediction in stock exchange markets. *In International Conference on Business Information Systems* (pp. 122-133). Springer, Cham.

Uysal, A. S., & Mulvey, J. M. (2021). A Machine Learning Approach in Regime-Switching Risk Parity Portfolios. *The Journal of Financial Data Science*, 3(2), 87-108.

Varenius, M. (2020). Using Hidden Markov Models to Beat OMXS30.

Vieriu, R. L., Goraş, B., & Goraş, L. (2011, June). On HMM static hand gesture recognition. *In ISSCS 2011-International Symposium on Signals, Circuits and Systems* (pp. 1-4). IEEE.

Visser, I., & Speekenbrink, M. (2010). depmixS4: an R package for hidden Markov models. *Journal of statistical Software*, 36(7), 1-21.

Waduge, N., & Ganegoda, U. (2018, December). Forecasting Stock Price of a Company Considering Macroeconomic Effect from News Events. *In 2018 3rd International Conference on Information Technology Research (ICITR)* (pp. 1-5). IEEE.

Waliszewski, K., & Warchlewska, A. (2021). Comparative analysis of Poland and selected countries in terms of household financial behaviour during the COVID-19 pandemic. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *16*(3), 577-615. https://doi.org/10.24136/eq.2021.021.

Zhang, J., Li, L., & Chen, W. (2021). Predicting stock price using two-stage machine learning techniques. *Computational Economics*, 57(4), 1237-1261.