

STOCK PRICE PREDICTION USING MARKOV CHAINS ANALYSIS WITH VARYING STATE SPACE ON DATA FROM THE CZECH REPUBLIC

Milan Svoboda¹, Pavla Říhová²

¹ University of West Bohemia in Pilsen, Faculty of Economics, Department of Economics and Quantitative Methods, Czech Republic, svobodm@kem.zcu.cz;

² University of West Bohemia in Pilsen, Faculty of Economics, Department of Economics and Quantitative Methods, Czech Republic, ORCID: 0000-0002-6632-445X, divisova@kem.zcu.cz.

Abstract: The article describes empirical research that deals with short-term stock price prediction. The aim of this study is to use this prediction to create successful business models. A business model that outperforms the stock market, represented by the Buy and Hold strategy, is considered to be successful. A stochastic model based on Markov chains analysis with varying state space is used for short-term stock price prediction. The varying state space is defined based on multiples of the moving standard deviation. A total of 80 state space models were calculated for the moving standard deviation with 5-step lengths from 10 to 30 in combination with the standard deviation multiples from 0.5 to 2.0 with the step of 0.1. The efficiency of the business models was verified for 3 long-term, liquid stocks of the Czech stock market, namely the stocks of KB, CEZ, and O2 within a 14-year period – from the beginning of 2006 to the end of 2019. Business models perform best when they use a state space defined on the length of a moving standard deviation between 15 and 30 in combination with multiples of the standard deviation between 1.1 and 1.2. Business models based on these parameters outperform the passive Buy and Hold strategy. In fact, they outperform the Buy and Hold strategy for both the entire period under review and the yielded five-year periods (including transaction fees). The only exception is the five-year periods covering 2015 for O2 stocks. After the end of the uncertainty period caused by unclear intentions of the new majority stockholder, the stock price rose sharply. These results are in conflict with the efficient markets theory and suggest that in the period under review, the Czech stock market was not effective in any form.

Keywords: Stock market prediction, technical analysis, Markov chains, efficient market theory.

JEL Classification: C02, C13, G14, G17.

APA Style Citation: Svoboda, M., & Říhová, P. (2021). Stock Price Prediction Using Markov Chains Analysis with Varying State Space on Data from the Czech Republic. *E&M Economics and Management*, 24(4), 142–155. <https://doi.org/10.15240/tul/001/2021-4-009>

Introduction

This empirical study deals with the short-term prediction of stock prices on the Czech stock market. Stock movements have been of interest to traders for a long time. Using a wide range of analytical methods, it tries to satisfactorily clarify past and present changes in stock prices. Based on these findings, it attempts to predict the future development of stock prices.

Early forecasting allows traders to make capital gains.

It is necessary to mention that according to Efficient Market Hypothesis (EMH), stock prices are unpredictable and markets are efficient. This means that the market responds immediately to any new information. This information cannot be predicted, it is randomly sent to the market and therefore the change in the exchange rate

is random and the exchange rates perform a so-called 'random walk'. In efficient markets the above-average profits cannot be achieved and according to this theory, other approaches are dysfunctional.

The idea of a random walk was probably first published in the doctoral dissertation of the French mathematician L. Bachelier in 1900. According to Cootner (1964), after more than 60 years this Bachelier study was first published in English and had recovered 'the rightful place' in financial econometrics. Later another scientist, Fama (1965, 1970, 1991), returned to this idea and developed it substantially. He examined the correlation between daily returns and prior's period daily returns for stocks composing the DJI index. The correlation coefficients were around zero. He thus concluded that the return on the stock had no relation to the return of the prior period. Fama's conclusions were confirmed by Solnik (1973) who worked with the data from European stock markets. The efficient market hypothesis (EMH) was also confirmed by other studies and researches by investment economists.

Roberts (1967) was the first to identify the particular forms of efficiency as a weak, semi-strong and a strong form of efficiency. The weak form of efficiency means that the exchange rate includes all information from historical data and therefore, methods of technical analysis are not able to predict the market. The semi-strong form of efficiency is a situation where the price includes both historical data and all public information and methods of fundamental analysis also fail here. The strong form EMH demonstrates the fact that even insider information is incorporated into the stock price. In a highly efficient market, therefore, the insider information is worthless and does not help investors to capture above-average returns.

In the beginning, EMH was widely accepted by the academic community. Later, however, studies questioning the theory started to emerge. For example, Shiller (1981) draws attention to the higher volatility of stock prices than can be explained by the volatility of dividends. Haugen (1999) believes that markets overreact to new unexpected information and refers to various studies that confirm the occurrence of anomalies in an efficient market.

The Czech stock market was also examined. In most cases, a weak form of efficiency is examined in the Czech stock market. The

works of Filáček et al. (1998) or Hanousek and Němeček (2001) inclined to the opinion that the Czech stock market is behaving inefficiently. On the contrary, more recent works by Filer and Hanousek (1996) or Diviš and Teplý (2005), tend to demonstrate that the stock market behaves efficiently in a weak form. Therefore, the methods of technical analysis do not bring above-average returns.

There is no consensus in academia assessing the question of whether or not markets are efficient. For example, research by Hájek (2007) pointed out that between 2000 and 2005, the Czech stock market was approaching a weak form of efficiency. Simultaneously, he claimed that the five-year period was too short. And that strategies that exploited of short-term dependences can be abnormally profitable. Research and discussions related to the efficiency of stock markets are likely to continue in the years to come. Current empirical research contributes to this discussion. This paper deals with a problem of predicting the direction of stock prices on the Czech stock market and builds on the work of Svoboda (2016). This work describes a stochastic model based on Markov chains analysis. The model is used to predict short-term future stock price movement.

The main aim of this empirical research is to modify the stochastic model mentioned above and apply this model successfully to the Czech stock market. It is possible to expect the following benefits of the research:

- Contribution to the discussion of whether the Czech stock market can be considered to have a weak form efficiency.
- Verification of the conclusions of the work by Svoboda (2016) in the longer term.

The paper is organized in the following way. The first part describes the theoretical background and the important theory on Markov chains. The second part of the work characterizes the data used to conduct research and describes the research methodology. The third most comprehensive part includes the obtained empirical results and their brief discussion. The final part contains the evaluation of the work performed and outline the direction of further research.

1. Theoretical Background

The research is based on the assumption that the stock price is continuously created on the stock exchange by the mutual interaction

between supply and demand. This supply and demand is generated by different types of traders (long-term investors, speculators). These traders have different time horizons, they use different methods to estimate the future income they will obtain by owning stocks, and they have different risk aversion, different amounts of capital. Every investor follows his intentions and has different reasons for buying or selling a given stock. Supply or demand for a given stock may not only be created due to the trader's subjective perception of the undervaluation or overvaluation of the stock, but also for many other reasons, such as the need to raise money, modification of investment strategy, raising new capital by secondary offering, forced purchases or sales in borrowed assets, repurchases of stocks by the joint stock company itself, etc. The simultaneous action of these factors results in the constant fluctuation of the stock price and can be seen as a random variable.

The next assumption of technical analysis is that stock prices fluctuate in trends. When the trend is identified in a timely manner, traders hope for above-average profits. The subject of this research is the forecasting of short-term price movements (tertiary trend), while it is not the price level that is important, but the estimation of the magnitude of price changes. The stock price fluctuates in short-term trends and during the duration of this trend the stock price accumulates a certain gain or loss against the price at the beginning of the trend. The greater this change, the greater the likelihood of a change in this trend. The key question for us is how large the accumulated loss or profit must be in order for the trend to change with a sufficiently high probability. According to the nature of the data (daily opening and closing prices), it is proper to use the theory of Markov chains to model the probability of a trend change.

Markov chains (MC) theory is described, for example, in Hillier and Lieberman (2010). MC are used for modelling processes which can be found in one of finite (countable) number of states in discrete time moments. MC is a sequence of discrete random variables x_1, x_2, x_3, \dots with the Markov property (at the time moment t_{n+1} the process will be in state i . It is stochastically dependent only on the state that was active at the previous time moment, i.e. at the time t_n). Formally it can be described as follows (1):

$$P(X_{n+1} = x_{n+1} | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n) \quad (1)$$

Particular realizations x_i are elements of a countable set $S = \{s_j\}, i = 1, 2, \dots, N$ which is called a state space. Behaviour of the described process is determined by:

- Vector of unconditional probabilities $p(n)^T = [p_1(n), p_2(n), \dots, p_N(n)]$, where T means transposition and $p_i(n)$ denotes probability that the process is in the moment n in the state i .
- Transition probability matrix P whose elements p_{ij} give conditional probability of process transition from the state i to the state j . That could be formally described as $p_{ij} = P(X_n = s_j | X_{n-1} = s_i)$, where $i = 1, 2, \dots, N$ and $j = 1, 2, \dots, N$.

If we know the probability of particular states appearance at the time moment when the process starts, we can describe the process behaviour using the relations (2):

$$p^T(n) = p^T(n-1)P \text{ thus } p^T(n) = p^T(0)P^n \quad (2)$$

It can be derived from the relations (2) that the long-term behaviour of this stochastic process is determined by the matrix P .

In order for the model to be successful, it is necessary to define the state space properly. There must be such states in the state space from which the process transforms into states with the opposite trend with a sufficiently high probability. In other words, we need to find states in which the trend is likely to change with a sufficiently high probability. Such states could be used to generate buying or selling orders.

The use of MC for modelling stock market behaviour is not a new issue. Over the last decade, a number of papers have been published using MC theory to model stock market development. In some of them the state space is defined very simply. For example, Doubleday and Esunge (2011) applied the Markov chains to the DJIA index and to a portfolio of selected stocks from this index. They defined a state space on daily price changes. They had tested two types of state space, with two states and with six states. Vasanthi et al. (2011) dealt with predicting the daily development of the value of stock indices. They predicted only the direction of price fluctuation, i.e. increase or decrease. In order to make the prediction they had used transition probabilities which they had

calculated successively from the last 5 years, 3 years and 1 year. Petković et al. (2018) used a three-state model for returns analysis on the Belgrade Stock Exchange. Similar methods have been applied to other exchanges, such as Lakshmi and Manoj (2020) used MC analysis for the Indian stock market or Yavuz (2019) for the Turkish stock market. The conclusions of these works show that such a simple definition of state space cannot be used to carry out successful business strategies and are now tested more often on emerging markets.

The most promising methods for successful trading seems to be the application of advanced techniques of MC analysis or a differently defined state space. Nguyen (2018) used the hidden Markov chains for monthly stock price prediction. He had tested models with two to six states. These models had been tested on S&P index. The results showed that the hidden Markov model have outperformed the B&H strategy and have yielded higher percentage returns. In the study by Huang et al. (2017) a model integrating two types of MCs was used for stock price prediction: regular and absorbing MC. The absorbing MC provides information on accumulated price changes during the length of tertiary trend. Riedlinger and Nicolau (2020) predicted the development of the FTSE 100 index one period ahead using multivariate MCs. The multivariate model enables to include more variables in the probability model. The authors assumed that stock prices could be affected by a number of variables and delays by more than one period.

The work by Svoboda (2016) dealt with the fact that stock prices may be retentive more than one period in such a way that the state space was defined on the basis of the size of daily cumulative changes in the stock price. The work shows that in such a defined state space there are states in which there is a sufficiently high probability of the tertiary trend reversal. The state space in the above-mentioned work is defined on the basis of multiples of the standard deviation from daily price changes. The standard deviation is calculated for the entire monitored period. However, the standard deviation calculated for the whole period does not respect one of the basic characteristics of stock markets that is changing volatility. Therefore, in this research, the state space definition is based on a moving standard deviation. It is expected that the state space determined in this way will be able to

respond to changing volatility and the model will demonstrate better results.

2. Data

We conduct research on stocks from the Czech stock market, which are traded on the Prague Stock Exchange (PSE). The criteria for including stock types in the research were as follows:

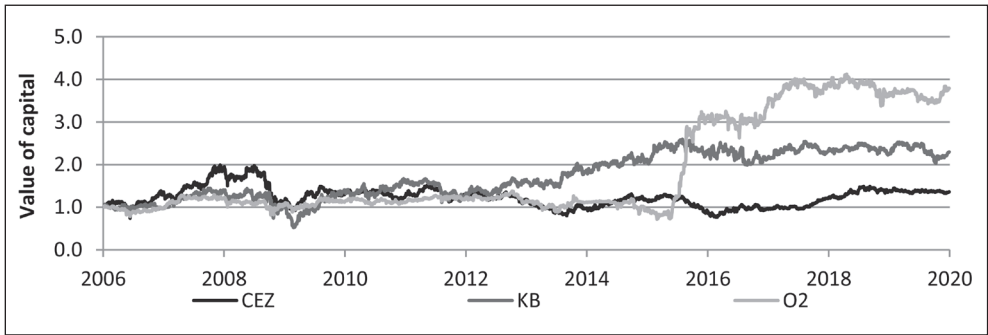
- PSE must be the main market for this stock type.
- The stocks must be traded on PSE long enough.
- Daily volumes must be high enough, i.e., fulfilling the assumption that the stock is traded by all types of investors and has sufficient liquidity.

These requirements are met only by the following few companies: the telecommunications company O2 CR (O2), the KB bank (KB) and the energy company CEZ (CEZ). We have daily opening and closing prices for each stock type for a fourteen-year period, from 2 January 2006 until 2 January 2020, i.e. approx. 3,500 business days. The data source is Patria Direct. The companies regularly paid dividends in the given period. In this research, the after-tax dividend is reinvested.

In the period under research, on 2016-05-12, KB split its stocks 5-for-1. All data (dividends, opening and closing prices) relating to KB were recalculated in order to maintain the data continuity. On 2015-06-01, O2 was divided into two companies: O2 and CETIN. The final price of the O2 stock the day before the division of the company was CZK 177.6. For one original O2 stock, the stockholder acquired one new O2 stock and one CETIN stock. On the first day after the split, the price of the CETIN stock was CZK 133.5 and the price of the O2 stock was CZK 69.2. At the time of the division of the company the market was informed that a majority stockholder of CETIN stocks plans to get rid of minority stockholders and the stocks will later be withdrawn from the market. Maintaining data continuity is addressed as follows. CETIN stocks are sold at the price of CZK 133.5 on the first day after the division, and this income is further counted in the same way as the dividend. This means that O2 stocks are purchased for them according to business strategy.

The process of the stock appreciation of the monitored stocks with the reinvestment of dividends is shown in Fig. 1.

Fig. 1: Development of capital value



Source: Patria, own research

Special attention in Fig. 1 is drawn to a sharp increase in the appreciation of O2 stocks in the second half of 2015. That increase was caused by the end of the uncertainty that had been evident as soon as the release of the information on the majority stockholder's actions entered the market. In autumn 2013, PPF purchased a majority stake from Telefónica.

3. Research Methodology

Stock prices must first be transformed into a suitable Markov chain. The transformation is performed in terms of an appropriately defined state space. After obtaining the MC, suitable states for generating trading signals are selected based on the transition probability matrix. These generated trading signals help to create trading strategies that serve as the footing for the trading model.

3.1 State Space

The definition of the state space is supported by the work Svoboda (2016). The stock space is defined on the cumulative changes in the stock price over the duration of the tertiary trend. The cumulative change in the stock price is denoted by Y_t , which is interpreted as short basic indices of daily closing prices. The base period is the day of the trend change, i.e., the transition point from decline to growth or vice versa. The duration of the trend is determined by the number of consecutive rising or falling closing prices. Formally, the calculation of the cumulative price change Y_t is described by (3):

$$Y_t = Y_{t-1} \frac{P_t}{P_{t-1}} \text{ if } (\dots \leq P_{t-2} \leq P_{t-1} \leq P_t) \quad (3)$$

$$\text{or } (\dots \geq P_{t-2} \geq P_{t-1} \geq P_t) \text{ else } Y_t = \frac{P_t}{P_{t-1}}$$

where P_t is the final daily price at time t , P_{t-1} is the final daily price at time $t - 1$ and P_{t-2} is the final daily price at time $t - 2$.

Based on the values of y_t (y_t is a percentage of Y_t) a state space is defined. A set with eight states are used to sort the data. As the stock price falls, the corresponding conditions are marked D_i . State D_1 is the state with the smallest price drop and state D_4 is the state with the highest price drop. Conditions when the stock price rises are marked G_i . The G_1 state is the state with the lowest price growth and, conversely, the G_4 state is the state with the highest price growth. The general state space model is defined by the following principle:

$D_4: y_t < -3\Delta_t$	$G_1: 0 \leq y_t < 1\Delta_t$
$D_3: -3\Delta_t \leq y_t < -2\Delta_t$	$G_2: 1\Delta_t \leq y_t < 2\Delta_t$
$D_2: -2\Delta_t \leq y_t < -1\Delta_t$	$G_3: 2\Delta_t \leq y_t < 3\Delta_t$
$D_1: -1\Delta_t \leq y_t < 0$	$G_4: 3\Delta_t \leq y_t$

where $\Delta_t = k\sigma_{t,l}$ represents the width of the interval, $\sigma_{t,l}$ is the standard deviation of daily changes in the stock price, k is the multiple of the standard deviation. The standard deviation is calculated as the moving standard deviation of length l , according to the formula (4):

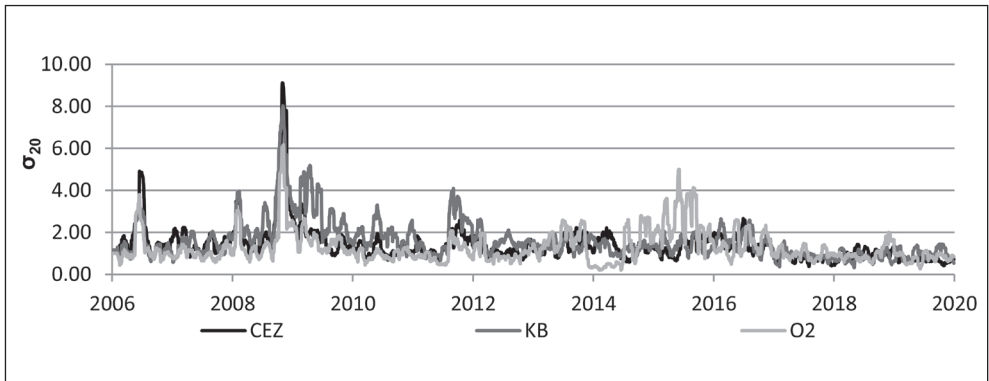
$$\sigma_{t,l} = \sqrt{\frac{1}{l} \sum_{i=0}^{l-1} (x_{t-i} - \bar{x}_{t,l})^2} \quad (4)$$

where x_{t-i} is the daily change of the stock price at day $t-i$, l is the length of the moving standard deviation (width of the sliding window) and $\bar{x}_{t,l}$ is the moving average lengths of l at day t . The so-defined state space model (varying model) can respond to changing volatility in stock markets. Svoboda (2016) in his work states, that space model (unvarying model) was defined on an unchanging standard deviation calculated from

all neighbouring values. For individual stocks, the fixed standard deviation is as follows: $\sigma_{O2} = 1.548$, $\sigma_{CEZ} = 1.663$, $\sigma_{KB} = 1.899$. The development of a moving standard deviation for parameter $l = 20$ is shown in Fig. 2.

The state assignment procedure for both models of the state space described above is illustrated on CEZ stocks with model parameters $k = 1$ and $l = 20$ in Tab. 1.

Fig. 2: Moving standard deviation for $l = 20$



Source: own

Tab. 1: Procedure for assigning states (CEZ, $k = 1$, $l = 20$)

Date	P_t	x_t [%]	y_t [%]	Unvarying model		Varying model	
				σ	State	σ_{20}	State
2019-12-12	505.5	0.999	0.999	1.66	G_1	0.71	G_2
2019-12-11	500.5	-1.476	-1.476	1.66	D_1	0.66	D_3
2019-12-10	508.0	0.594	1.195	1.66	G_1	0.59	G_3
2019-12-09	505.0	0.198	0.598	1.66	G_1	0.58	G_2
2019-12-06	504.0	0.099	0.398	1.66	G_1	0.57	G_1
2019-12-05	503.5	0.299	0.299	1.66	G_1	0.59	G_1
2019-12-04	502.0	-0.100	-2.240	1.66	D_2	0.60	D_4
2019-12-03	502.5	-1.374	-2.142	1.66	D_2	0.68	D_4
2019-12-02	509.5	-0.779	-0.779	1.66	D_1	0.63	D_2
2019-11-29	513.5	0.588	0.588	1.66	G_1	0.62	G_1
2019-11-28	510.5	-0.293	-0.293	1.66	D_1	0.61	D_1
2019-11-27	512.0	0.999	0.999	1.66	G_1	0.71	G_2
...	...						

Source: own

3.2 Transition Probability Matrix

Now that the MC is available it is proposed to find the probabilities of the transition between individual states. Before the calculation was performed MC filtration was carried out. This filtering procedure is meant to reveal the consecutive identical states. By filtration, we are able to skip the states in which the stock price stagnates (respectively, they change little in the same trend). These states are not considered to be of interest in terms of trading. For illustration we show a part of the string before filtering: ... D₄, G₁, G₁, G₁, G₂, D₁, D₁, D₂, G₂, D₂, D₂, ... and after filtration ... D₄, G₁, G₂, D₁, D₂, G₂, D₂, ... The matrix of transition probabilities **P** is determined for the filtered MC. The parameters $k = 1.0$ and $l = 20$ of the presented transition probabilities for both varying model and unvarying model are demonstrated in Tab. 2. These are average conditional transition probabilities calculated according to the formula (5):

$$p_{ij} = \frac{1}{3} (p_{ij}^{CEZ} + p_{ij}^{KB} + p_{ij}^{O2}) \quad (5)$$

where p_{ij}^{stock} represents conditional probability of transition between individual states. The ΣD_i and ΣG_i columns indicate the likelihood of staying in the trend or changing the trend. The n_i denotes the average number of occurrences of individual states is indicated.

From the values calculated in Tab. 2 we can observe:

- For both models, the probability of trend change in individual states is high enough, the unvarying model has a slightly higher probability.
- The values calculated by us for the unvarying model are almost identical to the values in Svoboda (2016), the increase in the time window did not bring about changes in the behaviour of stock prices.
- The anomaly can be noticed at the transition from G₄ to G₃ in the varying model, where zero would be expected. There has been a rare case where a moving standard deviation has risen sufficiently to be attributed to a lower G₁ state despite the rise in the stock price.

Tab. 2: Transition probabilities for the model for $k = 1.0, l = 20$

<i>t</i>		<i>t</i> + 1	D ₄	D ₃	D ₂	D ₁	G ₁	G ₂	G ₃	G ₄	ΣD _i	ΣG _i	<i>n</i>
D ₄	Unvarying	0	0	0	0	0.450	0.334	0.123	0.093	0	1	74.0	
	Varying	0	0	0	0	0.538	0.312	0.138	0.012	0	1	112.7	
D ₃	Unvarying	0.333	0	0	0	0.463	0.163	0.033	0.008	0.333	0.667	132.0	
	Varying	0.356	0	0	0	0.431	0.180	0.030	0.004	0.356	0.644	185.0	
D ₂	Unvarying	0.063	0.266	0	0	0.524	0.114	0.024	0.009	0.329	0.671	308.0	
	Varying	0.093	0.269	0	0	0.430	0.178	0.026	0.004	0.362	0.638	381.7	
D ₁	Unvarying	0.006	0.042	0.260	0	0.598	0.079	0.010	0.005	0.308	0.692	689.0	
	Varying	0.013	0.074	0.297	0	0.452	0.137	0.022	0.005	0.384	0.616	602.0	
G ₁	Unvarying	0.005	0.013	0.076	0.594	0	0.265	0.039	0.009	0.687	0.313	668.7	
	Varying	0.005	0.029	0.122	0.451	0	0.303	0.077	0.012	0.608	0.392	576.7	
G ₂	Unvarying	0.002	0.014	0.107	0.520	0	0	0.273	0.084	0.643	0.357	312.0	
	Varying	0.001	0.023	0.151	0.435	0	0	0.282	0.109	0.610	0.390	392.3	
G ₃	Unvarying	0.004	0.012	0.137	0.496	0	0	0	0.351	0.649	0.351	138.7	
	Varying	0.002	0.025	0.168	0.402	0	0	0	0.404	0.596	0.404	200.3	
G ₄	Unvarying	0.014	0.070	0.281	0.635	0	0	0	0	1.000	0.000	95.0	
	Varying	0.000	0.046	0.293	0.658	0	0	0.002	0	0.998	0.002	136.7	

Source: own

3.3 Trading Strategies and Models

In the following part, the term trading strategy identifies a set of rules that determine the purchase and sale of stocks. The trading model includes several trading strategies applied to selected stocks.

Trading strategies are based on the following principle. Purchasing signals are gradually generated by states D_3, D_4 and sales signals are generated by states G_3, G_4 . By combining buying and selling stock states, 4 trading strategies are obtained: $D_3-G_3, D_3-G_4, D_4-G_3, D_4-G_4$. Strategies D_3-G_3 imply that state D_3 generates purchase orders (if we no longer hold stocks) and state G_3 generates a sales order (if we hold stocks). This is in line with the conclusions of the work Svoboda (2016), where these 4 trading strategies generated the highest revenues.

Trading is always carried out according to the following rules:

- One trade (transaction) represents the purchase and subsequent sale of a stock.
- It is not possible to make two purchases in a row.
- If a buying or a selling signal is generated on a given day, the trade is executed with the opening price from the following day.
- The capital is always fully invested, so it is theoretically possible to buy parts of the stocks.
- Transaction fees are not taken into consideration.
- Dividends and other income (hereinafter referred to as dividends), if we are entitled to them, are reinvested after tax.
- A short selling (speculation on price decrease) is not taken into account.

The value of the invested capital is calculated according to the following formula (6):

$$C_n = C_0 \prod_{i=1}^n \frac{S_i + d_i}{B_i} \quad (6)$$

where $C_0 = 1.00$ is the initial value of capital (a unit of capital is invested), C_n is the value of capital after the n -th transaction, S_i is the selling price in i -th transaction, d_i represents the net dividends (and other income) in case of dividend day occurred during the i -th transaction, B_i is the purchase price in the i -th transaction. If the transaction fees were to be calculated it is sufficient to modify formula (6) into the following formula (7):

$$C_n = C_0 \left(\prod_{i=1}^n \frac{S_i + d_i}{B_i} \right) \left(\frac{1 - \frac{p}{100}}{1 + \frac{p}{100}} \right)^n \quad (7)$$

where p stands for the size of the fee (in percent).

It would be very risky for a trader to invest everything in one selected stock and use one trading strategy. To minimize the risk, it is appropriate to distribute capital evenly between the analysed stocks and between individual trading strategies. Therefore, individual trading strategies will not be evaluated as it is more significant to evaluate the entire portfolio of trading strategies, which is called the trading model. The trading model includes four trading strategies, defined on the same state space, which are applied to three stocks, i.e., the invested capital is divided into 12 equal parts. The evaluation of the trading model is represented by C_M and is formally determined by formula (8):

$$C_M = \frac{C_{O2} + C_{CEZ} + C_{KB}}{3}, \text{ where} \quad (8)$$

$$C_{stock} = \frac{C_{D3-G3} + C_{D3-G4} + C_{D4-G3} + C_{D4-G4}}{4}$$

According to EMH, trading in stocks following through with the above principles should not be successful and the value should not outperform the market. In other words, a passive stockholding should yield the same or a higher return. Passive stockholding can be referred to as a Buy and Hold (B&H) strategy. We will consider a trading model successful if it outperforms the B&H strategy. In this paper, the B&H strategy envisages the reinvestment of dividends. It is assumed that in the beginning of 2006 stocks were purchased and then held until the end of 2019. For the dividends paid, the stocks are purchased at current prices. By implementing the B&H strategy, the resulting value of capital for individual stocks would be $C_{CEZ} = 1.35$; $C_{KB} = 2.28$; $C_{O2} = 3.79$ so that the average value of capital is $C_{B\&H} = 2.47$.

4. Research Results and Discussion

In the first place, the summary results for individual models will be presented and then one of the models will be analysed in detail. The summary results are listed in Tab. 3. For each model, the value of C_M capital and the average number of realized trades n are given. A total of 80 varying state space models were analysed.

Tab. 3: Final value of capital

k	$l = 10$		$l = 15$		$l = 20$		$l = 25$		$l = 30$		Unvarying	
	C_M	n	C_M	n	C_M	n	C_M	n	C_M	n	C_M	n
0.50	1.46	200.0	1.50	187.5	1.34	185.5	1.54	180.4	1.70	180.5	1.72	141.8
0.60	1.66	179.4	1.83	168.0	1.85	166.0	1.99	164.7	2.01	162.8	2.24	122.3
0.70	2.05	155.4	2.19	152.1	1.98	145.5	1.96	143.3	2.21	142.5	2.30	105.5
0.80	2.02	134.3	2.32	129.4	2.28	128.0	2.18	123.5	1.84	119.8	2.11	86.3
0.90	2.10	115.1	2.50	111.1	2.27	108.7	2.32	106.9	2.09	104.7	2.38	74.3
1.00	2.31	98.5	2.32	93.4	2.73	92.7	2.18	91.4	2.04	88.3	2.56	63.0
1.10	2.54	83.8	3.04	81.0	2.80	78.3	2.51	75.8	3.37	75.6	2.63	51.9
1.20	2.63	69.7	2.71	68.3	2.93	65.5	2.83	63.9	2.91	63.9	2.56	43.8
1.30	2.67	61.1	2.37	56.7	2.54	54.7	2.42	54.2	2.22	53.1	2.40	36.6
1.40	2.47	52.0	2.23	47.6	2.51	47.4	2.27	46.1	2.29	44.9	2.67	31.5
1.50	2.51	43.3	1.95	37.8	2.18	39.5	2.19	38.8	2.05	38.4	2.36	26.0
1.60	2.31	36.8	2.09	33.8	2.22	34.1	2.04	33.0	2.00	32.7	2.22	22.8
1.70	2.16	31.3	2.31	30.1	2.39	29.5	2.47	28.7	1.88	28.2	1.81	19.8
1.80	2.20	26.4	2.17	25.2	2.29	25.1	2.30	23.8	2.35	25.1	1.83	17.6
1.90	2.15	23.3	1.79	20.2	2.00	20.1	2.10	21.2	2.12	21.1	1.96	16.0
2.00	1.99	19.7	1.70	16.8	1.82	16.6	2.01	18.0	2.64	18.0	1.98	14.8
Mean	2.20	83.1	2.19	78.7	2.26	77.3	2.21	75.8	2.23	75.0	2.23	54.6
St. dev.	0.323		0.375		0.392		0.281		0.411		0.295	

Source: own

The parameter k gradually took values from 0.5 to 2.0 with the step of 0.1. The parameter l for the length of the sliding standard deviation took values from 10 to 30 with the step of 5. An unvarying model was also calculated for each value of the parameters.

Results in Tab. 3 show that on average for all parameters k there is no significant difference between the unvarying and varying model. The average value of capital varies in a narrow range between 2.19 and 2.26. The best results are obtained when the parameter of the model k is between 1.1 and 1.4. While the unvarying model has a maximum capital value of 2.67 ($k = 1.4$), the varying models (except $l = 10$) achieve higher values: 3.04 ($k = 1.1$, $l = 15$), 2.93 ($k = 1.2$, $l = 15$), 2.83 ($k = 1.2$, $l = 25$), 3.37 ($k = 1.1$, $l = 30$). Varying models combining the parameters $l = 15$, $l = 20$, $l = 25$ and $l = 30$ with the parameters $k = 1.1$ and $k = 1.2$ slightly outperform the unvarying model and outperformed the passive B&H strategy

(2.47). As expected, it is also evident that the number of executed trades decreases with increasing k parameter. This is important when calculating trading fees.

The Tab. 3 shows the summary results for the entire 14-year period. Results may be skewed in one successful period, while in other periods, trading models may not be successful. In order to be considered reliable, trading models should perform well for any given period. The five-year sliding yields in Tab. 4 and five-year sliding annual yields in Tab. 5 are calculated for the model with the parameters $k = 1.2$ and $l = 25$. This model belongs to the group of models that demonstrated the best results. It has average results in this group. The length of five years was chosen because the minimum recommended investment horizon for investing in stocks is five years. In addition to the total returns of the trading model, the returns of individual stocks are also monitored and compared with the returns of the B&H strategy.

Tab. 4: Five-year yield for the model with parameters $k = 1.2$ and $l = 25$

Years	KB		CEZ		O2		C _M	
	Model	B&H	Model	B&H	Model	B&H	Model	B&H
2015-19	30.4%	7.3%	26.0%	15.3%	14.1%	301.2%	23.5%	107.9%
2014-18	65.0%	20.0%	54.6%	41.3%	-10.4%	227.0%	36.4%	96.1%
2013-17	88.1%	41.1%	42.5%	2.1%	-5.0%	238.8%	41.9%	94.0%
2012-16	155.0%	65.8%	26.8%	-25.8%	-9.2%	182.1%	57.6%	74.0%
2011-15	134.1%	39.3%	1.1%	-28.5%	-0.4%	178.8%	44.9%	63.2%
2010-14	174.0%	54.7%	32.7%	-12.0%	-5.3%	-17.5%	67.1%	8.4%
2009-13	159.5%	83.1%	23.7%	-15.2%	30.5%	7.7%	71.2%	25.2%
2008-12	121.6%	18.5%	2.3%	-38.3%	31.6%	-4.8%	51.8%	-8.2%
2007-11	82.8%	34.1%	28.7%	-3.7%	47.4%	23.2%	53.0%	17.9%
2006-10	84.5%	61.5%	41.6%	26.1%	53.7%	12.0%	59.9%	33.2%
2006-19	378.7%	128.3%	111.5%	34.8%	60.1%	279.2%	183.4%	147.4%

Source: own

Tab. 5: Average annual yield for the model with parameters $k = 1.2$ and $l = 25$

Years	KB		CEZ		O2		C _M	
	Model	B&H	Model	B&H	Model	B&H	Model	B&H
2015-19	5.5%	1.4%	4.7%	2.9%	2.7%	32.0%	4.3%	15.8%
2014-18	10.5%	3.7%	9.1%	7.2%	-2.2%	26.7%	6.4%	14.4%
2013-17	13.5%	7.1%	7.3%	0.4%	-1.0%	27.6%	7.2%	14.2%
2012-16	20.6%	10.6%	4.9%	-5.8%	-1.9%	23.1%	9.5%	11.7%
2011-15	18.5%	6.9%	0.2%	-6.5%	-0.1%	22.8%	7.7%	10.3%
2010-14	22.3%	9.1%	5.8%	-2.5%	-1.1%	-3.8%	10.8%	1.6%
2009-13	21.0%	12.9%	4.3%	-3.2%	5.5%	1.5%	11.4%	4.6%
2008-12	17.3%	3.5%	0.5%	-9.2%	5.6%	-1.0%	8.7%	-1.7%
2007-11	12.8%	6.0%	5.2%	-0.8%	8.1%	4.3%	8.9%	3.3%
2006-10	13.0%	10.1%	7.2%	4.7%	9.0%	2.3%	9.8%	5.9%
2006-19	11.8%	6.1%	5.5%	2.2%	3.4%	10.0%	7.7%	6.7%

Source: own

From Tab. 4 and 5 we see that the analysed model for KB and CEZ shares exceeds the B&H strategy for each five-year period. The average annual revenue for the entire 14-year period is 11.8% for KB (6.1% for B&H) and 5.5% for CEZ (B&H 2.2%). The average annual yield in five-year sliding windows ranged from 5.5% to 22.3% for KB (B&H from 1.4% to 12.9%), and for CEZ from 0.2% to 9.1% (B&H from -9.2% to

7.2%). We consider the differences between the model and the B&H strategy to be significant.

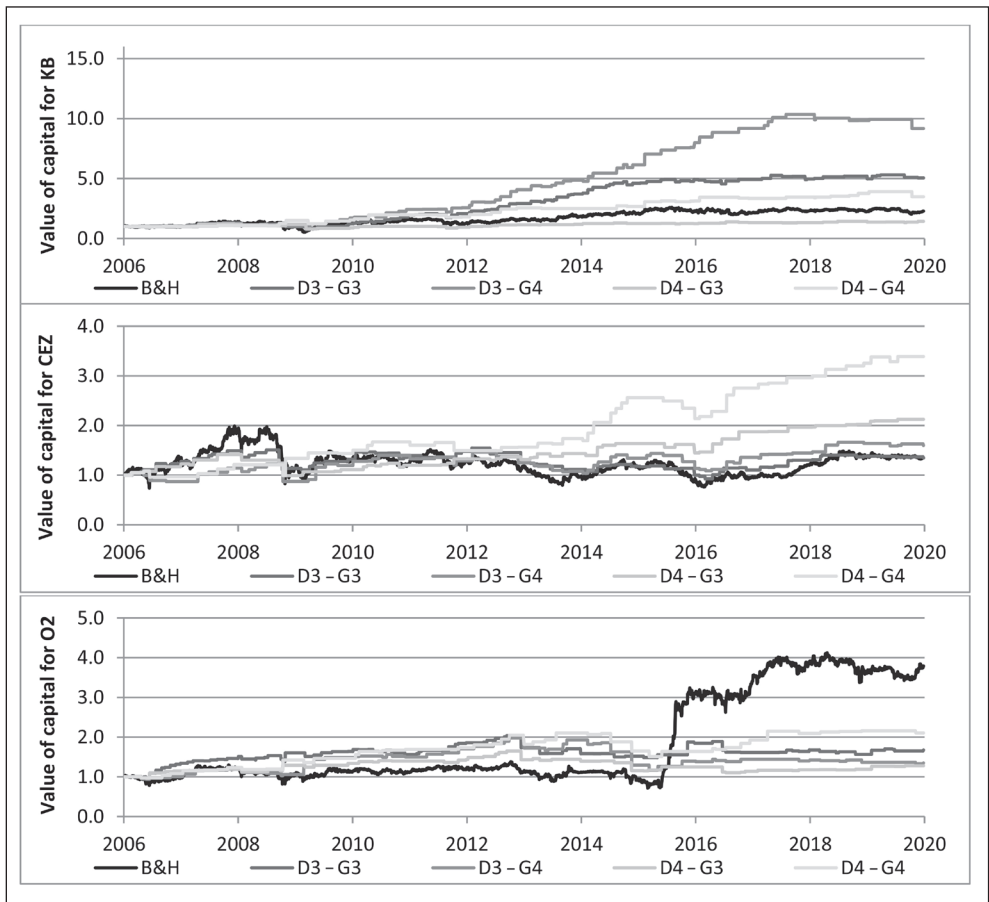
The average annual return on O2 shares for the entire 14-year period is 3.4% (B&H 10.0%). Throughout the period, B&H's passive strategy surpasses the business model. As mentioned earlier, 2015 can be considered completely atypical. This year brought out the uncertainty about the intentions of the new

Tab. 6: Value of capital for the model with parameters $k = 1.2$ and $l = 25$

Strategy	KB		CEZ		O2	
	C	n	C	n	C	n
D ₃ -G ₃	5.053	99	1.347	86	1.675	87
D ₃ -G ₄	9.175	71	1.597	65	1.349	58
D ₄ -G ₃	1.434	52	2.125	56	1.275	54
D ₄ -G ₄	3.484	45	3.392	49	2.102	44

Source: own

Fig. 3-5: Development of capital trading strategies for individual stocks



Source: own

majority owner, the public takeover bid and also the subsequent reduction of the free float to approx. 10% of the previous approx. 30%. In the independently evaluated period 2006–2014, the model achieved an average annual yield of 3.8% (B&H -0.6%) and the model surpassed the B&H strategy in each five-year period. In the four-year period 2016–2019, the model achieved an average annual yield of 1.1% (B&H 4.9%). The B&H strategy was never exceeded by the model in any given year from the given period. One possible explanation is that the reduction in free float resulted in much lower volumes of daily trades and short-term traders (speculators) may have ceased to be interested in O2 shares.

For practical trading, trading fees must also be taken into account. Each broker has a different amount of fees. Fees are usually constructed as a certain percentage of the trade volume plus possibly some fixed payment. In general, the higher the trade volume, the lower the percentage of the fee. The lowest fees at CZK 50,000 start at 0.0015% of the trade volume. In the analysed model, an average of 63.9 trades took place in 14 years with an average of 4.6 trades per year. Substituting into formula (7) at a fee of 0.0015% of the volume of the trade and taking the considered 4.6 trades per year, we can observe that the value of capital is reduced by approximately 1.3% per year. The business model surpasses B&H strategy even when the fees are included.

The following concluding analysis will address the contribution of individual trading strategies to the overall revenue of the model. Again, the model with parameters $k = 1.2$ and $l = 25$ will be analysed. The final value of capital and the number of realized trades for individual trading strategies is given in Tab. 6.

The data in Tab. 6 confirms, as mentioned earlier, that with increasing D (buy signal is generated with a larger price decrease) and with increasing G (sell signal is generated with a larger price increase), the number of executed trades decreases. It is demonstrated that the returns of individual strategies that are applied to the same stock can vary significantly. From the results in Tab. 6 it is also clear that it is important to use a combination of multiple trading strategies in order to reduce risk. The development of capital trading strategies for individual trading strategies is shown in Fig. 3–5.

It is possible to make the following conclusions based on Fig. 3–5:

- There is no single strategy (not even within a single stock type) that would perform better than the rest of the strategies for the whole period. This can be clearly seen in the D3–G4 strategy for KB stocks. Until 2012, the performance of this strategy was comparable to other strategies. From 2012 to about 2017, this strategy achieved exceptionally high returns. Since 2018, this strategy has been creating losses. We believe that one of the factors that influences the success of individual strategies might be the volatility of stock prices, which changes over time.
- None of the strategies is able to respond to a situation in which the share price is rising sharply. For CEZ shares, we can observe this in the period from mid-June 2006, when the price was CZK 540 until December 2007, when the share price was above CZK 1,400. This can be seen in O2 shares in 2015, when at the beginning of June 2015 the share price was CZK 70 and in December 2015 the price was CZK 250.

Conclusions

This empirical study focuses on short-term stock price predictions based on a stochastic Markov chain (MC) analysis model. This study was started in the work of Svoboda (2016). The aims of the study were to verify the methods described in this work over a longer period of fourteen years; to modify the described methods and to contribute to the discussion on the effectiveness of Czech stock market. Predictive models are considered successful if trading based on these models outperforms passive buy and hold (B&H) strategy.

In the original work, the MC state space was defined on the basis of a cumulative change in the share price and a fixed standard deviation; in this study, the fixed standard deviation was replaced by a sliding standard deviation. Empirical results have shown that both approaches can be considered successful and outperform the B&H strategy. Business strategies based on sliding standard deviation models achieve slightly better results. Strategies are also successful when fees are included.

The study also indicated possible weaknesses where strategies fail. Strategies

have not been successful in situations in which the share price rises sharply. Likewise, the strategies were unsuccessful in case the daily trading volumes decreased due to a reduction in the free float of O2 shares. One possible explanation is that speculators may have ceased to be interested in these shares.

The results of this study suggest that the Czech stock market is not behaving effectively. The business models for KB and CEZ shares surpassed the B&H strategy in every five-year sliding window, even when calculating fees. For O2 shares, the models were successful only until 2014, after 2014 the models were not successful. However, the fact that the business model was not successful does not mean that stock prices behaved effectively in the given period, it indicated that one specific trading method might not have worked for them.

Therefore, throughout the fourteen-year period, buy and sell orders were generated on the basis of only one factor, namely accumulated price changes, and these results may be considered to be very good and are worth examining further. Possible directions for further research are to:

- Identify which events trigger specific stock behaviours and select the most effective trading strategies based on the expected stock behaviours. This means that the business model would use various business strategies at different times.
- Combine Markov chains studies with other tools of fundamental or technical analysis.
- Confirm the results of this work by including other shares in the portfolios, verify whether the results depend on the volume of traded shares and extend the time period.
- Apply these models to foreign stock markets.

Acknowledgments: This research was supported by Grant SGS-2018-042 of the Faculty of Economics, the University of West Bohemia in Pilsen.

References

Cootner, P. H. (Eds). (1964). *The Random Character of Stock Market Prices*. Cambridge, MA: MIT Press.

Diviš, K., & Teplý, P. (2005). Information efficiency of Central Europe stock exchanges (in Czech). *Czech Journal of Economics and Finance (Finance a úvěr)*, 55(9–10), 471–482.

Doubleday, K. J., & Esunge, J. N. (2011). Application of Markov Chains to Stock Trends. *Journal of Mathematics and Statistics*, 7(2), 103–106. <https://doi.org/10.3844/jmssp.2011.103.106>

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, 38(1), 34–105. <https://doi.org/10.1086/294743>

Fama, E. F., (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383–417. <https://doi.org/10.2307/2325486>

Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5), 1575–1617. <https://doi.org/10.2307/2328565>

Filáček, J., Kapička, M., & Vošvrda, M. (1998). Testování hypotézy efektivního trhu na BCPP [Efficiency market hypothesis: Testing on the Czech capital market]. *Czech Journal of Economics and Finance (Finance a úvěr)*, 48(9), 554–566.

Filer, R. K., & Hanousek, J. (1996). *The Extent of Efficiency in Central European Equity Markets* (CERGE-EI Working Paper Series No. 104). Prague: Center for Economic Research and Graduate Education – Economics Institute. <https://doi.org/10.2139/ssrn.1542343>

Hájek, J. (2007). Czech Capital Market Weak-Form Efficiency, Selected Issues. *Prague Economic Papers*, 16(4), 303–318. <https://doi.org/10.18267/j.pep.310>

Hanousek, J., & Němeček, L. (2001). Czech parallel capital markets: discrepancies and inefficiencies. *Applied Financial Economics*, 11(1), 45–55. <https://doi.org/10.1080/09603100150210255>

Haugen, R. A. (1999). *The New Finance: The Case Against Efficient Markets* (2nd ed.). Hoboken, NJ: Prentice Hall.

Hillier, F. S., & Lieberman, G. J. (2010). *Introduction to Operations Research* (9th ed.). New York, NY: McGraw-Hill.

Huang, J.-C., Huang, W.-T., Chu, P.-T., Lee, W.-Y., Pai, H.-P., Chuang, C.-C., & Wu, Y.-W. (2017). Applying a Markov chain for the stock pricing of a novel forecasting model. *Communications in Statistics – Theory and Methods*, 46(9), 4388–4402. <https://doi.org/10.1080/03610926.2015.1083108>

Lakshmi, G., & Manoj, J. (2020). Application of Markov process for prediction of stock market performance. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(6), 1516–1519. <https://doi.org/10.35940/ijrte.F7784.038620>

- Nguyen, N. (2018). Hidden Markov Model for Stock Trading. *International Journal of Financial Studies*, 6(2), 36. <https://doi.org/10.3390/ijfs6020036>
- Petković, N., Božinović, M., & Stojanović, S. (2018). Portfolio optimization by applying Markov chains. *Anali Ekonomskog fakulteta u Subotici*, 54(40), 21–32. <https://doi.org/10.5937/AnEkSub1840021P>
- Riedlinger, F. I., & Nicolau, J. (2020). The Profitability in the FTSE 100 Index: A New Markov Chain Approach. *Asia-Pacific Financial Markets*, 27(1), 61–81. <https://doi.org/10.1007/s10690-019-09282-4>
- Roberts, H. W. (1967). *Statistical versus Clinical Prediction of the Stock Market CRSP* (Unpublished manuscript). Chicago, IL: University of Chicago.
- Shiller, R. J. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *The American Economic Review*, 71(3), 421–436.
- Solnik, B. H. (1973). Note on the Validity of the Random Walk for European Stock Prices. *The Journal of Finance*, 28(5), 1151–1159. <https://doi.org/10.1111/j.1540-6261.1973.tb01447.x>
- Svoboda, M. (2016). Stochastic Model of Short-term Prediction of Stock Prices and Its Profitability in the Czech Stock Market. *E&M Economics and Management*, 19(2), 188–200. <https://doi.org/10.15240/tul/001/2016-2-013>
- Vasanthi, S., Subha, M. V., & Nambi, S. T. (2011). An empirical study on stock index trend prediction using Markov chain analysis. *Journal of Banking Financial Services and Insurance Research*, 1(1), 72–91.
- Yavuz, M. (2019). A Markov chain analysis for BIST participation index. *Journal of Balikesir University Institute of Science and Technology*, 21(1), 1–8. <https://doi.org/10.25092/baunfbed.433310>