VILNIUS UNIVERSITY

JOANA KATINA

# INVESTIGATION OF PREDICTION PROBLEMS BY THE VIRTUAL STOCK EXCHANGE

Summary of Doctoral Dissertation

Technological Sciences, Informatics Engineering (07 T)

Vilnius, 2015

Doctoral dissertation was prepared at the Institute of Mathematics and Informatics of Vilnius University in 2009–2014.

**Scientific Supervisor –**

Prof. Habil. Dr. Jonas Mockus (Vilnius University, Technological Sciences, Informatics Engineering – 07 T).

**The dissertation will be defended at the Council of the Scientific Field of Informatics Engineering of Vilnius University:**

**Chairman –**

Prof. Dr. Habil. Albertas Čaplinskas (Vilnius University, Technological Sciences, Informatics Engineering – 07 T).

**Members:**

Prof. Dr. Andrej Brodnik (University of Ljubljana, Physical Sciences, Informatics – 09 P),
Assoc. Prof. Dr. Olga Kurasova (Vilnius University, Physical Sciences, Informatics Engineering – 07 T),
Doc. Dr. Saulius Minkevičius (Vilnius University, Technological Sciences, Informatics – 09 P),
Prof. Dr. Habil. Edmundas Kazimieras Zavadskas (Vilnius Gediminas Technical University, Technological Sciences, Informatics Engineering – 07 T).

The dissertation will be defended at the public meeting of the Council of the Scientific Field of Informatics Engineering in the auditorium number 203 at the Institute of Mathematics and Informatics of Vilnius University, at 10 a. m. on the 7th of September, 2015.

Address: Akademijos st. 4, LT-08663 Vilnius, Lithuania.

The summary of the doctoral dissertation was distributed on the 7th of August, 2015.
A copy of the doctoral dissertation is available for review at the Library of Vilnius University or on this website: *www.vu.lt/lt/naujienos/ivykiu-kalendorius*

VILNIAUS UNIVERSITETAS

JOANA KATINA

# PROGNOZAVIMO PROBLEMŲ TYRIMAS VIRTUALIOJE AKCIJŲ BIRŽOJE

Daktaro disertacijos santrauka

Technologijos mokslai, informatikos inžinerija (07 T)

Vilnius, 2015

Disertacija rengta 2009–2014 metais Vilniaus universiteto Matematikos ir informatikos institute.

**Mokslinis vadovas –**
prof. habil. dr. Jonas Mockus (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – 07 T).

**Disertacija ginama Vilniaus universiteto Informatikos inžinerijos mokslo krypties taryboje:**

**Pirmininkas –**
prof. habil. dr. Albertas Čaplinskas (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – 07 T).

**Nariai:**
prof. dr. Andrej Brodnik (Liublijanos universitetas, fiziniai mokslai, informatika – 09 P),
doc. dr. Olga Kurasova (Vilniaus universitetas, fiziniai mokslai, informatikos inžinerija – 07 T),
doc. dr. Saulius Minkevičius (Vilniaus universitetas, technologijos mokslai, informatika – 09 P),
prof. habil. dr. Edmundas Kazimieras Zavadskas (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija – 07 T).

Disertacija bus ginama Vilniaus universiteto viešame Informatikos inžinerijos mokslo krypties tarybos posėdyje 2015 m. rugsėjo mėn. 7 d. 10 val. Vilniaus universiteto Matematikos ir informatikos instituto 203 auditorijoje.

Adresas: Akademijos g. 4, LT-08663 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2015 m. rugpjūčio mėn. 7 d.
Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: *www.vu.lt/lt/naujienos/ivykiu-kalendorius*

## 1. Introduction

### *Research area and relevance of the problem*

There is a lot of research in the field of forecasting with different forecasting models and methods: AR($p$) – Auto-Regression, ARMA($p$,$q$) – Auto-Regression Moving Average, ARIMA($p$,$d$,$q$) – Auto-Regression Integrated Moving Average, ARFIMA($p$,$d$,$q$) – Auto-Regression Fractionally Integrated Moving Average, ARCH – Auto-Regression Conditional Heteroskedasticity, ANN – Artificial Neural Networks, etc. More complex models provide a better fitting (less prediction errors in respect of past data) to historical data, as usual. The problem is how these models will predict future financial data, since the economic conditions are changing.

A popular opinion is that, although the good past behavior does not guarantee a good future outcome. One can expect at least a positive correlation. However, the recent results show that, under some unfavorable conditions, the correlation is negative. Even under favorable conditions, one needs very long time series for a positive correlation. This effect is described by the general term over-fitting, borrowed from the ANN field.

The emphasis on the price prediction methods is justified assuming that more accurate price predictions provide better profits. In mathematical terms, it means that profit is an increasing function of prediction accuracy of asset prices. However, the assumption is not necessarily true, especially in the problem of portfolio optimization which involves many stocks and is using complicated investment strategies in the presence of different constraints. The investment strategy is defined as a pair of the trading rule (buy-sell conditions) and the prediction model. Under some conditions, the positive correlation of portfolio profits and the price prediction errors happens to be possible. It means negative relations of portfolio profits and the accuracy of price predictions.

This possibility is illustrated by some experiments of this work. It is the natural result of such prediction models, where greater differences between the present and past stock prices generate larger deviations of the predicted prices from the present ones. It can produce both the positive and negative components of the correlation between price changes and price prediction errors. For example, if the positive components of correlations prevail, then it is a source of occasional positive correlations of portfolio profits and price prediction errors, since the profits are on average increasing functions of the price changes.

An additional source of this counter-intuitive phenomenon may be that in the multi-stock case, the profit is defined for the whole portfolio, but not for individual stocks while the prediction errors are calculated separately for each stock. The profit depends on two factors: the prediction model and the trading rule. The experiments of this work indicate that the trading rule is the most important factor.

When investigating the relation of optimal profits and the prediction accuracy, an additional difficulty is that the profit optimization has to be performed in the space of investment strategies. It can be reduced to maximization of an increasing function of differences between the present and past asset prices with many constraints, describing the conditions of an individual investor. The traditional prediction models minimize squared or absolute deviations of predicted prices from the actual ones, as usual. Therefore, the minimal prediction error does not necessarily provide maximal profits, since different optimization criteria are involved.

The description and theoretical explanation of this apparently new result was not found in publications. Investigations of the relation between functions that define the portfolio profit and prediction errors are an important topic for further theoretical and experimental investigation. The main objective is to define such conditions where the portfolio profit is a monotonous function of the prediction accuracy. This work was designed to investigate both the influence of over-fitting and the relation of profits to the prediction errors. The second problem is, apparently, more important, since it is almost an open area yet, in the field of financial optimization.

To reduce the influence of environmental and temporary factors, a virtual stock exchange model was developed. This model simulates interactions of major investors, using different investment strategies, with a view to maximize profits. The influence of a large number of smaller investors is represented as the Gaussian noise. To observe the over-fitting directly, the historical data of four time periods, representing different economic conditions, were investigated. To estimate the profit-prediction error relation, the respective correlations were calculated, using both the real and virtual data. In this work, the correlations were defined between the profit and prediction error instead of the prediction accuracy. Therefore the positive correlation means the negative relation of profit and prediction accuracy.

The best investment strategies differ in different time periods and in real and virtual markets. As usual, simpler prediction models provide smaller prediction errors. It seems natural in the context of over-fitting and corresponds to the efficient market hypothesis.

The new and the most important result of this work is the establishment of the fact that minimal prediction errors do not necessarily provide maximal profits. In both the real and virtual markets, positive profit-prediction error correlations happen under some conditions. An important part of future investigation belongs to the virtual stock exchange since this model by definition represents just some basic rules of the game what can be controlled by researchers and does not depend on the changing environment.

### *The research object*

The research object here is a virtual stock market and its application by defining the statistical correlation between the prediction error and actual profit, using different stock trading rules in real and virtual markets.

### *The aim and tasks of the research*

The aim of this work is to improve the previously developed stock exchange model, which simulates stock exchange processes, and to apply it to prediction problems with a view to establish a relation between customers' profit and stock price prediction accuracy in real and virtual stock markets.

***The research methods***

Information search, classification, analysis, benchmarking, and generalization methods were used for analyzing theoretical and experimental results in the optimal finance investment (portfolio) problem, including predictions and stock exchange models. The next day price was predicted using autoregressive prediction methods. Based on the experimental research method, the analysis of statistical data and research results was made. The benchmarking method was used to summarize all these results.

***Scientific novelty***

The main scientific novelty of the individual part of this work is application of a stock exchange model (virtual stock exchange) in the prediction problem and explanation why the correlation between portfolio profit and prediction error is insignificant statistically and could be negative. It means that higher prediction accuracy does not ensure greater profit and customers that use less accurate predictions, could get higher profits in particular situations.

***The practical value of the study results***

The conclusion that the prediction accuracy alone cannot ensure higher profit is important for the future investment optimization research. A particular attention should be paid to direct search of robust trading rules, less sensitive to unpredictable market changes. An improved stock exchange model could be used for developing new other stock exchange models.

***Statements presented for defense***

Using different trading rules, the correlation between the prediction accuracy and actual profit is statistically insignificant and, in some cases, negative both in real and virtual stock markets. It means that actual profit obtained using realistic trading rules, is not a monotonous prediction accuracy function.

***Approbation of the research results***

The main results of the thesis were published in 5 articles: 2 in peer-reviewed scientific publications, 2 in other scientific publications, 1 in conference proceedings. The main results were also presented and discussed at 8 international and national conferences in Lithuania.

***The structure and scope of the dissertation***

The thesis work consists of five chapters, the list of references, and appendices. The titles of the thesis chapters are: Introduction, Research review, PORTFOLIO model, Experimental research, and Conclusions. Tables, illustrations, and a list of used markers and abbreviations have also been provided in the thesis. The total scope of the work without the appendices is 115 pages that include 86 figures, 58 formulas, and 3 tables. The list of references consists of 60 sources.

## 2. Research review

### *The preceding works and a Stock exchange game model*

A simple Stock exchange game model was introduced in (Mockus, 2003) to simulate the behavior of several stockholders using fixed buying-selling margins at fixed bank yield. In (Mockus, 2010; Mockus & Raudys, 2010; Mockus, 2012), the model was investigated and compared with real data. In (Mockus, 2010; Mockus, 2012), it has been proposed to apply the Nash Equilibrium (NE) to strategies that define buying-selling margins and bank haircuts dynamically. That enables us to simulate market illiquidity that is an important feature of the present financial crisis (Allen, 2008).

In (Mockus, et al., 2014), the initial version of the PORTFOLIO model is introduced. The model includes the transaction costs to reflect the reality better. To represent users that prefer linear utility functions, the AR-ABS($p$) autoregressive model, minimizing the absolute values, is added to the traditional AR($p$) model, minimizing the least squares. In this paper, the new version of the PORTFOLIO model is described and investigated in the context of over-fitting and profit-prediction error relation.

The objective of PORTFOLIO is not forecasting, but simulation of financial time series that are affected by predictions of the participants. The “virtual” stock exchange can help in testing the assumption of rational investor’s behavior vs. the recent theories that explain financial markets by irrational responses of major market participants (Krugman, 2000; Krugman, 2008; Krugman, 2009). The model has been compared with actual financial time series detecting both the similarities and differences.

The model is designed as a tool to represent the behavior of an individual investor who wants to predict how the expected profit depends on different trading rules, using different forecasting methods of real and virtual stocks. It is assumed that the available information is the previous stock rates and only the mean values of assets are predicted.

The purpose of the model is to explore the relationship between the real data and theoretical models, between profits and prediction errors, and to investigate what other results can be obtained using this simple model.

### *Optimal financial investment problem*

The optimal financial investment (Portfolio) problem, including the forecasting and market models, was investigated by the leading financial organizations and hundreds of scientists. This problem is also topical for small investors who want to invest their own capital, to save or enlarge it. A special attention is paid to the financial market analysis and, as an evidence of it, to a considerable number of Nobel Prizes and various publications. Therefore, only the most influential publications are mentioned in this work: (Bernanke, 2004; Greenspan, 2009; Krugman, 2012; Krugman, 1996; Krugman, 2009; Krugman, 2008; Sharpe, 2000; Sharpe, 1964; Sharpe, 1963; Scholes, 1998) (Fama, 1970; Ait-Sahalia & Hansen, 2009; Hansen & Sargent, 2007; Hansen & Scheinkman, 2012; Hansen, 2013; Hansen & Scheinkman, 2012; Hansen, 2012; Anderson, et al., 2012; Hansen, 2012; Arellano, et al., 2012; Hansen & Sargent, 2011; Borovička, et al., 2011; Akerlof & Shiller, 2009; Shiller, 2008; Shiller, 2009; Shiller, 1998; Shiller, 1989; Shiller, 2007; Merton, 1971; Markowitz, 1959) (Markowitz, 1952; Sharpe, 1994; Sharpe, 1966; Merton, 1972; Black & Scholes, 1973; Nash, 1951; Bagočius, et al., 2014; Dadelo, et al., 2014). The citations include those authors to who were awarded the Nobel Prize in Economics for the problems related to the theory of financial markets. The aims of most of these works are forecasting, portfolio optimization, risk minimization, and capital distribution. In the financial market research, the market prediction and portfolio optimization were often regarded together. However, in most of the financial market investigations, forecast and investment problems were carried out separately.

Some works, close to the results of this paper, are also included. For example, (Bagočius, et al., 2014; Dadelo, et al., 2014) presents examples of multi-criteria optimization that will be applied in the future developments of this work. In this paper, all the investors apply the same single criteria, namely, the expected profit.

Also, an important part of the financial market analysis is the behavior of market participants. There are different assumptions on this issue: some scientists say that it is rational, and others that it is irrational. It is a very important point, because it can explain many processes of financial market. The aim of the works (Merton, 1971; Markowitz, 1959; Markowitz, 1952; Sharpe, 1994; Sharpe, 1966; Merton, 1972; Black & Scholes, 1973; Nash, 1951; Sharpe, 2000; Sharpe, 1964; Sharpe, 1963; Scholes, 1998; Fama, 1970) was to provide theory, recommendations and tools for diversification of the assets, depending on the acceptable risk level, assuming the rational behavior of participants. In the works (Krugman, 2012; Krugman, 1996; Krugman, 2009; Krugman, 2008; Hansen, 2013; Hansen & Scheinkman, 2012; Hansen, 2012; Anderson, et al., 2012; Hansen, 2012; Arellano, et al., 2012; Hansen & Scheinkman, 2012; Hansen & Sargent, 2011; Borovička, et al., 2011; Akerlof & Shiller, 2009; Shiller, 2008; Shiller, 2009; Shiller, 1998; Shiller, 1989; Shiller, 2007; Merton, 1971), irrational behavior was considered as an important factor. In the works (Ait-Sahalia & Hansen, 2009; Hansen & Sargent, 2007; Hansen & Scheinkman, 2012), some robust statistical techniques for analysis and forecasting of economic data were developed.

The contribution of the recent Nobel Prize winners (Fama, 1970; Hansen, 2013; Hansen & Scheinkman, 2012) may be summarized in the following way: Eugene Fama and several collaborators have demonstrated that stock prices are extremely difficult to predict in the short run, and that new information is very quickly incorporated into prices. These findings changed the market practice. Robert Shiller discovered in the early 1980 that stock prices fluctuate much more than corporate dividends, and that the ratio of prices to dividends tends to fall when it is high, and to increase when it is low. Lars Peter Hansen has developed a statistical method that is particularly well suited to testing rational theories of asset pricing. The laureates have laid the foundation for the current understanding of asset prices, which relies in part on fluctuations in risk and risk attitudes, and partly on behavioral biases and market frictions.

The problem of optimal investment was discussed in many well-known publications. Recently, the practical impotence of the problem has been increasing when major investors were accompanied by millions of small stockholders without financial education and experience deciding how to use their savings better.

The traditional approach is represented by the Modern Portfolio Theory (MPT). The Nobel prizes were awarded to MPT author Harry Markowitz and to William Forsyth Sharpe, who developed methods of reflecting the investment risk (Sharpe Ratio). The development of these theories and investigation of alternative approaches are described in major financial publications and discussed at international conferences.

A number of investment funds are making decisions by the software, based on the theoretical results of Robert C. Merton and Myron S. Scholes. Some limitations of this theory were noticed during the recent financial crisis when the investors have experienced considerable losses.

An effective approach to financial market investigation is the creation of its model. There are many types of models that simulate the financial (stock) market or its part: stock market games, market simulators, forecast models, tools for market process analysis, and others.

Financial market simulators are developed to satisfy the needs of small individual investors. The examples are the StockTrak global portfolio simulator and MarketWatch, virtual stock exchange. Some banks offer their own investment simulators such as the Barclay's Fantasy Investment Game. The users of these simulators working with "Virtual Stocks" are informed about the results. The graphical user interfaces are friendly. However, the theoretical base of these models and computing algorithms remains unknown. So the users cannot grasp the reasons why they win and why they experience losses.

The models of financial markets were investigated, assuming random interactions of independent financial agents. For example, in the paper (Darley & Outkin, 2007), an artificial stock market is considered by learning agents.

### *On features of the stock exchange model*

The default PORTFOLIO model of financial market includes a number of assets and stockholders. In addition to time series of real assets, the time series of virtual financial market are generated by simulating the interactions of different investors, using their own individual investment strategies. In the basic version, a number of different investment strategies are presented, including the ones based on the Modern Portfolio Theory. The individual preferences are regarded, using the utility theory and several heuristic procedures, developed in cooperation with real stockholders. The means are provided for creating additional investment strategies by users themselves.

The simulated investment procedures include autoregressive prediction models. The models that minimize the mean absolute error (MAE) are added to the traditional ones minimizing the mean squared error (MSE). The results of the virtual financial market are compared with historical data of different real assets.

The purpose of PORTFOLIO is meant to explore the relationship between the real and virtual data, between the profits and statistical prediction error, and to investigate what other results can be obtained using this simple model. The interactive mode of buying-selling levels allows manipulation of virtual financial markets. The scientific objective of this approach is to test the hypothesis that some important stock exchange features can be approximately described as a game of players using rational strategies.

In the virtual stock exchange, it is supposed that stock prices are primarily the result of the game of several major stockholders with some random deviations that represent large numbers of small investors. Investment decisions depend on stockholders' predictions of the future stock prices and expected dividends.

PORTFOLIO assumes that each player predicts stock prices by autoregressive models AR($p$) or AR-ABS($p$) of order $p$. The parameters of the model AR($p$) are estimated using the standard least squares algorithm for different $p$. The parameters of the AR-ABS($p$) model, which represents the linear utility function, are defined by the linear programming.

Actual stockholders use their own ways of predicting. The AR($p$) and AR-ABS($p$) models are as the simplest initial approximations of prediction processes. Simulations using AR($p$) models with various $p$ indicate that the prediction errors do not deviate significantly from the simplest models with $\boldsymbol{p = 1}$ representing the basic assumptions of the efficient market theory. However, the profits, simulated by PORTFOLIO depend, on $p$ significantly.

In contrast to the well-known published results, this work not only simulates the traditional results of utility and portfolio theories, but also complements them by various investment procedures and presents a possibility to users to develop and implement their own investment strategies. That is important for individual users with different approaches to risk.

The model can be useful for financial studies, scientific collaboration, and to stockholders which are investigating optimal investment problems and regard risk in the individual way. The new elements introduced in the paper, such as application of individual utility functions defining the personal portfolio, are helpful.

The risk evaluation by individual utility functions is basically different from the traditional evaluation by the return variance. The reason is that individual investors are risk-prone regarding small sums and risk averse while investing large sums. In this work, special global optimization procedures are used to solve this problem. It is not the domain of MPT. In this work, the special global optimization procedures are applied to the task. In addition, the algorithms and software for testing the personal utility functions were created (Mockus, 2014). These new tools are intended to help individual stockholders to select investment strategies by their personal preferences.

In short, the specific features of the proposed model are as follows:

1. Optimization in the space of investment strategies instead of portfolio optimization.
2. Implementation of both the real and virtual stock market in a single model.
3. Possibility to analyze the results (price prediction errors and profits) of 190 different investment strategies, including ten trading rules and nineteen forecasting models using real and virtual data.
4. Possibility to optimize the investment, assessing the risk by less restricted individual utility functions.

## 3. PORTFOLIO model

***Trading strategies, single stock market***

The aim of this work is the investigation of the weak and often negative relation between portfolio profits and accuracy of predictions of asset prices first observed in (Mockus, et al., 2014), and confirmed by additional experiments of the present paper. To explain this unexpected phenomenon, one needs the details of the stock exchange model. The model is the result of joint work. The dissertation (Katin, 2014) was aimed to establish a relationship between the virtual and real stock market. The aim of this work is to use a stock market model to determine the relationship between the portfolio profits and stock price prediction accuracy.

To explain the new results, a description of the stock-market model, used in this work, is required. Therefore, the description of the main parts of the model is continued, including some improvements and additional details.

In the beginning, a formal description, considering single stock trading by $I$ major players using the simplest rules, was given. The notation and basic expressions of the later PORTFOLIO version (Mockus, et al., 2014) were kept and supplemented with the new ones of this updated model.

The following symbols are used to denote the basic values, parameters, and variables:

$z(t) = z(t, i)$ is the price at time $t$, predicted by the player $i$,

$Z(t)$ is the actual[1]price at time $t$,

$U(t) = U(t, i)$ is the actual profit accumulated at time $t$ by the player $i$,

$\delta(t)$ is a dividend at time $t$,

$\alpha(t)$ is a yield at time $t$,

$\gamma(t)$ is the interest rate at time $t$,

$\beta(t, i)$ is a relative stock price change at time $t$ as predicted by the player $i$:

$$\beta(t, i) = \frac{z(t+1, i) - Z(t)}{Z(t)}. \tag{1}$$

In the PORTFOLIO model, the investor’s decisions depend on the expected profitability[2] (relative profit). It is defined as the relative profit $p(t, i)$ of an investment at

---

[1] The term “actual” means simulated by PORTFOLIO.

time $t$. The relative profit $p(t,i)$ depends on the predicted change of stock price $\beta_i(t)$, dividends $\delta_i(t)$, the yield $\alpha(t)$, and the interest $\gamma(t)$

$$p(t,i) = \begin{cases} \beta(t) + \delta(t) - \gamma(t), & \text{investing borrowed money,} \\ \beta(t) + \delta(t) - \alpha(t), & \text{investing own money.} \end{cases} \tag{2}$$

The aim is profit, thus a customer $i$:

a) will buy some number $n_b(t,i) \geq n(t)$ of stocks, if profitability is greater comparing with the relative transaction cost $\tau(t,n)$; $p(t,i) > \tau(t,n)$,
b) will sell stocks, if the relative loss (negative profitability $-p(t,i)$) is greater as compared with the transaction cost $p(t,i) < -\tau(t,n)$,
c) will do nothing, if $-\tau(t,n) \leq p(t,i) \leq \tau(t,n)$. Here the relative transaction cost is defined as the relation:

$$\tau(t,n) = \frac{\tau_0}{n(t)Z(t)}, \tag{3}$$

where $\tau_0$ is the actual transaction cost and $n = n(t)$ is the number of transaction stocks. From the equality $\tau(t,n) = p(t,i)$ it follows that the minimal number of stocks to cover transaction expenses is

$$n(t) = \frac{\tau_0}{p(t,i)Z(t)}. \tag{4}$$

Therefore, the buying-selling strategy $S(t,i)$ of the customer $i$ at time $t$ in terms of profitability levels is

$$S(t,i) = \\ = \begin{cases} \text{buy } n_b(t,i) \geq n(t) \text{ stocks}, & \text{if } p(t,i) \geq \tau(t,n) \text{ and } n \leq n_b^{\max}, \\ \text{sell } n_s(t,i) \geq n(t) \text{ stocks}, & \text{if } p(t,i) \leq -\tau(t,n) \text{ and } n \leq n_s^{\max}, \\ \text{wait}, & \text{if } |p(t,i)| \leq \tau(t,n^{\max}), \end{cases} \tag{5}$$

where $n^{\max} = \max(n_b^{\max}, n_s^{\max})$, where $n_b^{\max}$ is the maximal number of stocks to buy, and $n_s^{\max}$ is the maximal number of stocks to sell.

*Gaussian Model to Represent Small Stockholders*

The actual price of a stock at time $t+1$ is defined as the price of a previous deal of major stockholders plus the noise $\varepsilon(t)$. The deal happens if the selling stockholder has stocks to sell and the buying stockholder has sufficient funds.

$$Z(t+1) = \begin{cases} z_b(t,n) + Z(t) + \epsilon(t+1), & \text{if } Z(t) < z_b(t,n), \\ z_s(t,n) + Z(t) + \epsilon(t+1), & \text{if } Z(t) > z_s(t,n), \\ Z(t) + \epsilon(t+1), & \text{if no deal.} \end{cases} \tag{6}$$

*Buying-selling price*

The market buying price at time $t$ is the largest buying price of players $i = 1, \ldots, I$:

$$z_b(t,n) = z_b(t,n,i^{\max}), \tag{7}$$

where $i^{\max} = arg \max_i z_b(t,n,i)$.

The market selling price at time $t$ is the lowest selling price of players $i = 1, \ldots, I$:

$$z_s(t,n) = z_s(t,n,i^{\min}), \tag{8}$$

---

[2] The term "profit" can define losses if negative terms prevail.

where $i^{\min} = arg \min_i z_s\,(t,n,i)$.

The number of stocks owned by the player $i$ at time $t+1$ is

$$N(t+1,i) = \begin{cases} N(t,i) + n_b(t,n,i), & \text{if } Z(t) < z_b(t,n), \\ N(t,i) - n_s(t,n,i), & \text{if } Z(t) > z_s(t,n), \\ N(t,i), & \text{if no deal}, \end{cases} \qquad (9)$$

where $n_b(t,n,i)$ and $n_s(t,i)$ are the numbers of stocks for buying and selling operations by the player $i$ at time $t$. In PORTFOLIO it is assumed, for simplicity, that the total number of stocks $N_{\text{sum}}$ is not limited.

*Investors' Profit*

The product $N(0,i)\,Z(0,i)$ is the initial investment to buy $N(0,i)$ shares by the investors' own capital at the initial price $Z(0,i)$. The initial funds to invest are $C_0(0,i)$ and the initial credit limit is $L(0,i)$.

$L(t,i)$, $t = 1,\dots,T$ is the credit available for a customer $i$ at time $t$. The investors' own funds $C_0(t,i)$ available for investing at time $t$ are defined by the following recurrent expression:

$$C_0(t,i) = C_0(t-1,i) - (N(t,i) - N(t-1,i))\,Z(t), \qquad (10)$$

where $t = 1,\dots,T$. Here the product $(N(t,i) - N(t-1,i))\,Z(t)$ defines the money involved in buying-selling stocks.

Stocks are obtained using both investor's own money $C_0(t,i)$ and the funds $b(t,i)$ borrowed at the moment $t$.

*Multi-Level Operations*

To represent risk-aware stockholders, one needs at least three buying profitability levels $p_b(t,i,l)$, $l = 1,2,3$, where

$$\begin{gathered} p_b(t,i,l+1) > p_b(t,i,l), \\ p_b(t,i,1) = \tau(t), \end{gathered} \qquad (11)$$

and three selling profitability levels $p_s(t,i,l)$, $l = 1,2,3$, where

$$\begin{gathered} p_s(t,i,l+1) < p_s(t,i,l), \\ p_s(t,i,1) = -\tau(t), \\ p_b(t,i,l) > p_s(t,i,1), \end{gathered} \qquad (12)$$

to explain the behavior of major stockholders. The level $l = 1$ means to buy-sell just one stock. The level $l = 3$ means to buy-sell as many stocks as possible, and the level $l = 2$ is in the middle.

***Trading rules***

*Trading rule No. 1, Risk-aware stockholders: buying the best – selling the losers by three profitability levels.*

Consider the operations, involving different stocks denoted by indexes $j = 1,\dots,J$. Denote by $p(t,i,j)$ the profitability of the $j$th stock for a customer $i$ at time $t$. Denote by $j^{\max}$ the stock with the highest profitability:

$$j^{\max} = \arg\max_j p\,(t,i,j). \qquad (13)$$

First, the stockholder $i$ sells all nonprofitable stocks:

$$p_s(t,i,j) \le -\tau(t,i,j), \qquad (14)$$

and then invests all the available funds to buy the most profitable stock. This selling strategy reflects risk-regarding users who keep some less profitable stocks to avoid possible losses, if the predictions happen to be wrong.

In the PORTFOLIO model, the number of sold stocks by a few major players is not equal to the total number of stocks bought by these players. The assumption is that the exact balance is provided by a large number of small stockholders which are buying, if the prices are low, and selling, if the prices are high.

*Trading rule No. 2, Risk-aware stockholders: buying the best – selling all the losers*

First, the stockholder $i$ sells all nonprofitable stocks:

$$p_s(t,i,j) \leq -\tau(t,i,j), \tag{15}$$

and then invests all available funds to buy the most profitable stocks. The stockholder $i$ does not sell the stock $j$, if the expected loss is smaller than the transaction cost $|p(t,i,j)| < \tau(t,i,j)$.

This selling strategy reflects risk-aware users who keep some less profitable stocks to avoid possible losses, if predictions happen to be wrong. Note that the risk-regarding users keep a part of losing stocks as well.

*Trading rule No. 3, Risk-neutral stockholders: buying the best – selling all the rest*

First, the risk-neutral stockholder is selling all the stocks except for the most profitable ones to raise funds for buying the single most profitable stock. Thus they maximize the expected profit and ignore the risk. That explains the term risk-neutral stockholders. Such behavior is rational for very rich users who invest just a small proportion of their wealth.

*Trading rule No. 4, Risk-averse stockholders: selling and buying in proportion to profitability*

Consider the operations, involving different stocks denoted by indexes $j = 1,\ldots,J$. Denote by $p(t,i,j)$ the profitability of the $j$th stock for a customer $i$ at time $t$. Denote by $J_+$ a set of stocks with the positive profitability and by $J_-$ the stocks with the negative profitability. Denote $J_b = |J_+|$ and $J_s = |J_-|$.

$$j_+^{\max} = \arg\max_{j \in J_+} p\,(t,i,j), \tag{16}$$

and

$$j_-^{\min} = \arg\min_{j \in J_-} p\,(t,i,j). \tag{17}$$

First, stocks are sold in proportion to negative profitability levels. Then all the accumulated resources are used to buy stocks in proportion to positive profitability levels. Thus, the risk-averse users are buying some stocks with less profitable predictions. In this way, they distribute the funds more equally as compared with other types of users and so are less prone to the risk factor.

*Trading rule No. 5, Defining risk by survival probabilities and the individual utility function*

An alternative to MPT is the risk evaluation by the individual utility functions that determine particular investors' profit-to-risk relation (Fishburn, 1964; Mockus, et al., 1997; Sharpe, 2007; Mockus, 2014). In (Sharpe, 2007), this problem was solved assuming that the marginal utility is an additive decreasing function. It is equivalent to assuming that the investor is risk-averse. It excludes the investors who behave like risk-prone while investing small sums and become risk-averse when large sums are involved. To represent these investors, the additive utility functions are considered, when the marginal utility is not necessarily decreasing. The advantage is the increased flexibility.

The difficulty is multimodal objective functions, involved in the portfolio optimization. The software test for automatic evaluation of a person's utility function is implemented in (Mockus, 2014).

Now an illustrative example describes here how to optimize the investment of some fixed capital in Certificates of Deposit (CD) and Stocks, using individual user defined utility functions.

Investing in CD, the yields $\alpha_i$ are defined by contracts. Only the reliability $p_i$, $i = 1, \dots, n_b$ of banks is uncertain. Investing in stocks, in addition to the reliability $p_i$, $i = n_b 1, \dots, n$ of companies, their future stock rates are uncertain, too. The predicted relative stock rates are defined by the coefficient $\beta_i$ that shows the relation between the present and predicted stock rates. The prediction "horizon" is supposed to be the same as the maturity time of CD.

To simplify the model, suppose that one predicts $L$ different values of relative stock rates $\beta_i^l$, $l = 1, \dots, L$ with the corresponding estimated probabilities $p_i^l$, $\sum_{l=1}^{L} p_i^l = 1$, $p_i^l \geq 0$. It means $L$ multi-valued coefficients $c_i^l = 1 + \beta_i^l$, $l = 1, \dots, L$ for stocks and single-valued coefficients $c_i = \beta_i$ for CD.

In this case, one can define probabilities $p(y^i)$ of discrete values of wealth $y^i$, $i = 1, \dots, n + m$ by exact expressions.

An advantage of this approach is flexibility and a good theoretical basis. A disadvantage is a large amount of calculations needed to maximize the utility function which can be multi-modal if this function is not convex. However, the main difficulty of this approach is a reliable definition of survival probabilities. Therefore, in the next section, a version of diversification, defined by maximizing some modification of the traditional Sharpe ratio, is implemented.

*Trading rule No. 6, Risk-avoiding users, maximizing the Sharpe ratio in the context of the modern portfolio theory (MPT)*

MPT is a mathematical formulation of diversification in investing, with the view of selecting a collection of investment assets that haves collectively a lower risk than any individual asset. The diversification lowers the risk even if the assets are positively correlated (Markowitz, 1959; Markowitz, 1952; Merton, 1972).

MPT models an asset return as a stochastic function and defines risk as the standard deviation of return. MPT defines a portfolio as a weighted combination of assets, so that the return of the portfolio is the weighted combination of asset returns. By defining the weights of different assets, MPT seeks to reduce the total variance of the portfolio return. A risk-free asset can be included in the portfolio, as well.

In (Sharpe, 1966), the Sharpe ratio is defined as:

$$S = \frac{E[R_a - R_b]}{\sigma} = \frac{E[R_a - R_b]}{\sqrt{var[R_a - R_b]}}, \tag{18}$$

where $R_a$ is the asset return, $R_b$ is the return on a benchmark asset, such as the risk-free rate of return or an index such as S&P 500. $E[R_a - R_b]$ is the expected value of excess of the asset return over the benchmark return, and $\sigma$ is the standard deviation of the expected excess return.

The data from time $t = 1$ to $t = T$ is the learning set. The testing set would be from $t = T + 1$ up to $t = 2T$. To simplify the expressions, one can assume that the available funds are $I(t) = 1$ with the corresponding adjustment of scales.

The cyclic processes in the world of finances are not considered in this work.

*Applying short-term rules in the long-term investment*

In the short-term trading, the previous data are used to predict the next-day stock rate. The investment decisions by four heuristic trading rules were based on these predictions.

Similar rules are used to develop four additional long-term rules. The difference is that the stock rates are predicted not daily but only once by whole the learning set, for example, three, six, or twelve months. Investments are made based on the predictions, and the assets are sold at the end of the testing period of the same duration as the learning period, as usual. These four trading rules are numbered from No. 7 to No. 10, respectively.

***Prediction models***

*AR(p) model*

Assume that the player *i* predicts the next-day stock prices $z(t+1, i)$ using the AR($p$) model (Cochrane, 1997). Professional investors are trying to obtain additional information on the fundamentals of the stock and use sophisticated statistical models. Thus, the AR($p$) model can be regarded as the simplest simulator of a nonprofessional player who is making investments, based on the data observed during the past *p* days.

The profit of the player *i* depends on the accuracy of prediction $\beta(s, i)$ made at time *s* ($s = 1, \ldots, t$, where *t* denotes the present time).

Assume that the stock rates change following the simple relations below:

$$Z(s+1) = \sum_{k=1}^{p} a_k Z(s-k+1) + \varepsilon_{s+1}. \tag{19}$$

This formula describes the traditional autoregressive model AR($p$) of order $p$. However, in the context of this paper, relation (19) reflects opinions of stockholders who are making investment decisions based on the optimal next day predictions obtained using the past data. An important task of this work is to investigate how the portfolio profits depend on the standard statistical prediction errors.

To represent risk-neutral users, the AR-ABS model was applied by minimizing the absolute residuals instead of the squared ones. The details of PORTFOLIO implementation of the prediction models are in (Mockus, 2012; Mockus, et al., 2014).

*Comparison of different ARIMA settings*

An important part of the expert system, intended for stock-exchange imitation, is the prediction models which are part of the general trading rules. Therefore, some experiments were done investigate how the prediction error depends on ARIMA parameters.

Fig. 1 shows how the averages of MAE and MSE depend on the ARIMA parameters $p$ and $q$, as the parameter $d = 0$.

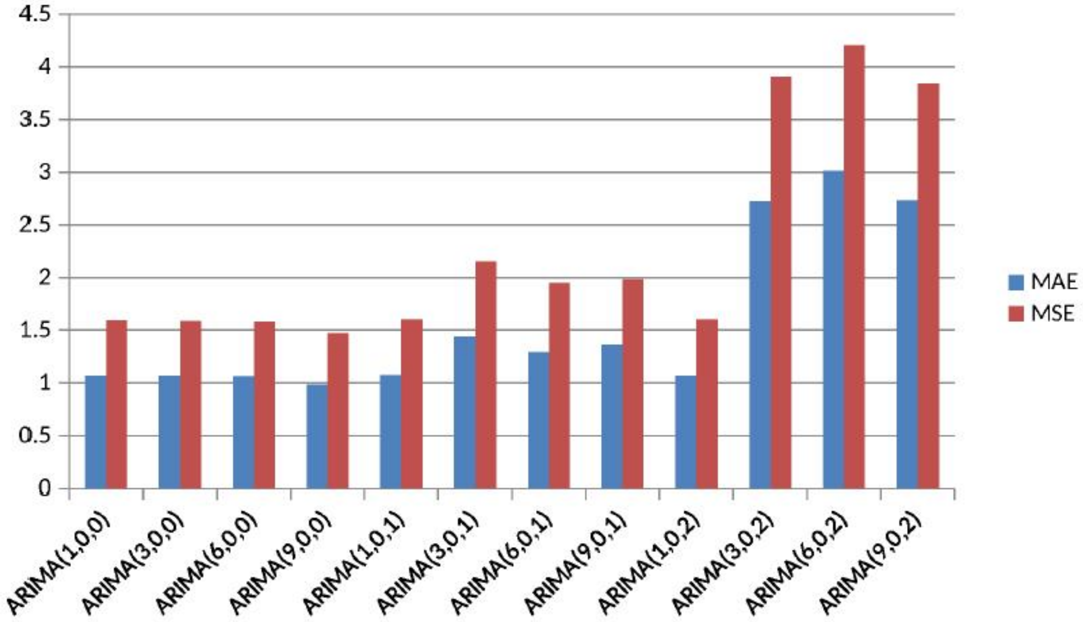


**Figure 1** *MAE and MSE as a function of the parameters p and q.*

The averages of MAE and SE are in USD, since the contribution of more expensive stocks into the profit is greater. The result is that the minimal MAE and SE were achieved by ARIMA($p$,0,0) that represents the autoregressive model AR($p$). It is in accordance with the recent results on over-fitting (Bailey, et al., 2014) and the traditional theory of Efficient Markets by (Fama, 1995) and is one of the reasons why, in most experiments, only autoregressive models were considered. Another reason is the computational complexity of ARIMA models involving the global optimization.

### *PORTFOLIO model for software implementation*

To run software on any device or computer, it should be written as a platform independent. Java is one of the best programming languages that could be used for the multiplatform application development. It is also good for calculations and computations. In 2015, Java is one of the most popular and most used programming languages by the TIOBE index. Therefore this language was chosen for realizing the PORTFOLIO model.

The prototype of the PORTFOLIO model was a part of the general online system for graduate studies and scientific collaboration (Mockus, 2006; Mockus, 2003). The code of the updated PORTFOLIO2 model is available in the Git version control system BitBucket. This command should be used to clone the program:

```
git clone https://bitbucket.org/igoriochek/stock_joana.git
```

That can be done both through the command line or using any IDE, for example, NetBeans. This code will create a Git local repository, in which the user could fully work with the project: read the program code, make updates and changes, run the program. A mathematical description is in (Mockus, et al., 2014).

The objective of the PORTFOLIO model is not forecasting, but simulation of stock exchange processes that are affected by predictions of the participants. Multi-stock extensions and a number of different trading rules represent both the heuristics of potential investors and the well-known theoretical investment strategies.

That makes the model more realistic and allows the portfolio optimization in the space of investment strategies, in both the historical and virtual environments. This is an essential improvement as compared with the traditional single-stock models with a direct interaction of investment agents.

The "virtual" stock exchange can help in testing the assumption of rational investor behavior vs. the recent theories that explain financial markets by irrational responses of major market participants.

The model is designed as a tool to represent the behavior of an individual investor, who wants to predict how the expected profit depends on different investment rules using different forecasting methods of real and virtual stocks. It is assumed that the only available information is the historic data of real stocks.

Optimization in the space of investment strategies and implementation of both the real and virtual stock market in the single model are the new properties of the PORTFOLIO model. An unexpected result was that the minimal stock price prediction errors do not necessarily provide the maximal profits. Therefore, the complete information is presented for independent testing and verification of this important new result. This result can be tested and verified independently without special skills and equipment. All the experimental conditions are defined and reproducible.

The profits of both investors and banks are calculated. The PORTFOLIO model simulates the behavior of a group of investors, who trade stocks in real and virtual environments. The optimization is performed on a set of investment strategies. The investor strategies include prediction models and trading rules. This is the main specific feature of the PORTFOLIO model.

Using this software, investors can choose one out of 190 investment strategies, including ten trading rules and nineteen forecasting models. Three of these trading rules simulate the known theoretical results, the rest are new and simulate heuristics of different investors with different approaches to risk.

When investigating the real environment, historical stock prices of popular international companies can be used. In the virtual environment, prices are generated by simulating the behavior of up to eight different major investors. The random noise simulates the influence of small investors.

The structure of software renders a possibility to extend the model: to add new prediction methods and trading rules. In large-scale automatic experiments, the MySQL technology was applied, using the NetBeans and XAMPP tools. So, the software can be used, modified, tested, and verified independently.

Fig. 2 shows the software operation scheme. It reflects the workflow of the model and shows the main components (large blocks).

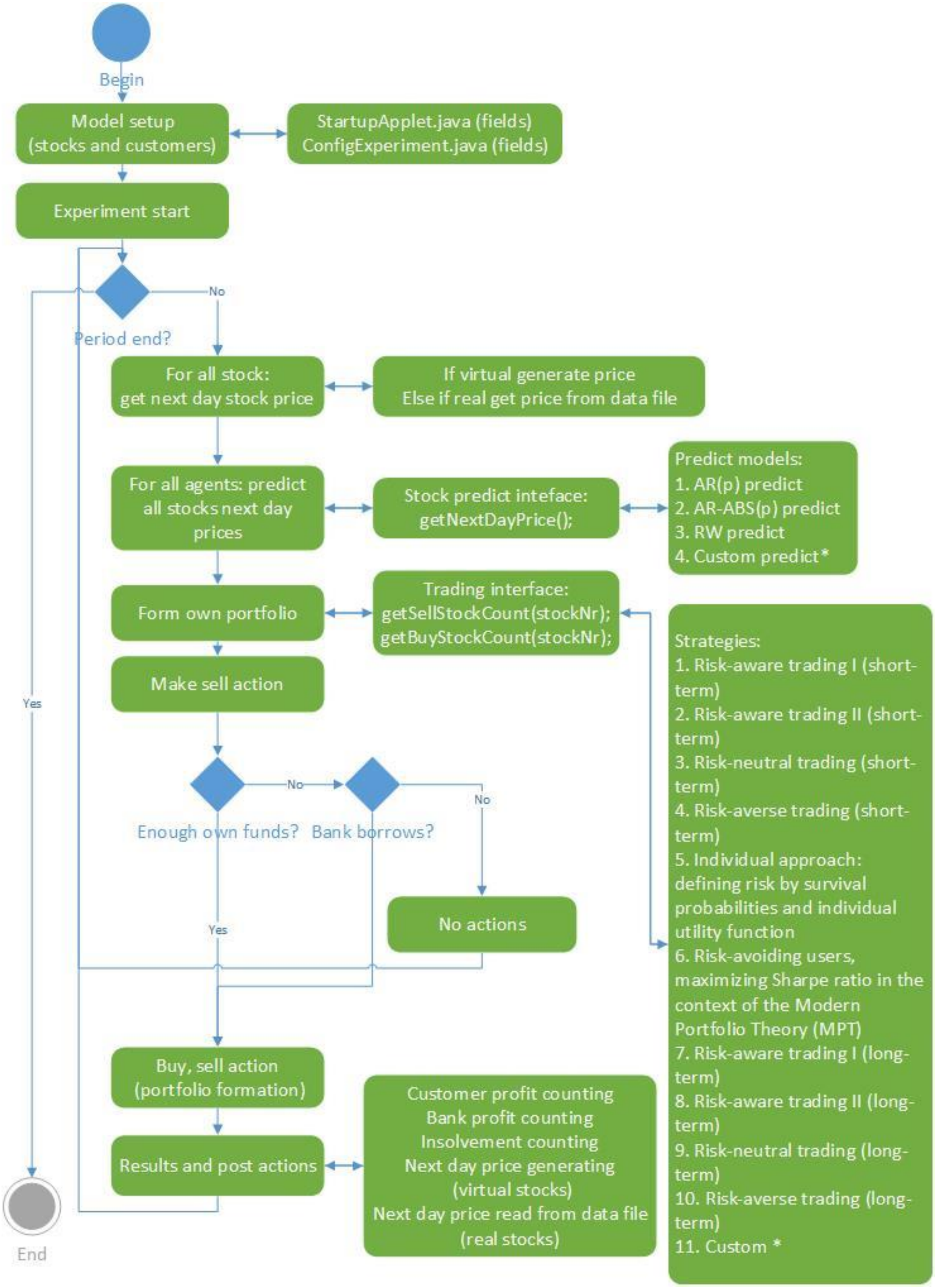


**Figure 2** *The software operation scheme*

## 4. Experimental research

In this section, the main experimental results of the new PORTFOLIO model are presented. Both real and virtual modes were investigated.

In the real mode, three sets of historical daily close prices were downloaded into PORTFOLIO directly by *http://finance.yahoo.com*. These included:

Period I. 364 working days from 2009-01-03, this is a period of economic growth after the crisis.

Period II. 364 working days from 2012-02-07, that is a newer, more stable time.

Period III. 352 working days from 2012-07-19.

Period IV. 364 working days from 2013-06-29, this period shows the present time.

The historical data of the following eight stocks were used: MSFT (Microsoft Corporation), AAPL (Apple Inc.), GOOG (Google Inc.), NOK (Nokia Corporation), TM (Toyota Motor Corporation), BAC (Bank of America Corporation), BA (The Boeing Company), ORCL (Oracle Corporation).

In the virtual mode, the stock prices were generated by simulating the buying-selling behavior of eight virtual investors. The initial prices were defined at the start of simulation, the next day prices were generated by the simulation. The average results of 100 independent samples were recorded.

Ten trading rules have been implemented, including four short-term rules and six longer term ones. Different rules represent a different approach to risk. The trading strategy is defined as a pair of the trading rule (TR) and the prediction model, used by the rule. Mainly the simple autoregressive models were regarded. They differ by the memory parameter p and the fitting criteria. Two criteria were considered: the traditional least squares one, denoted by AR, and the minimum of absolute deviations, denoted by AR-ABS. In addition, the Random Walk (RW) model was implemented, too.

Setting $p$ values from 1 to 9, eighteen different prediction models can be generated. The 19th model is RW. Thus, 190 different trading strategies were implemented. The model is open and renders a possibility to extend the list of strategies by implementing additional trading rules and prediction models. In this paper, a subset of 80 trading strategies (selected from the set of 190) has been investigated.

Sometime specific details will be presented, describing the Period I and the virtual market. The growth tendency is a common feature of both the virtual market and the Period I of real time. The contribution of more stable Periods II and III will be reflected, describing the average results of all the three periods of time.

### *Real stock experiment*

In this section, the experiment with historical data of the Period I is discussed. This period represents the recovery of economic and financial activities in the post-crisis time.

Table 1 shows the best portfolios, defined using ten trading rules and eight prediction models. The most profitable portfolio was provided by the strategy TR1, AR(6). It contains mainly BAC stocks. These stocks are also the main part of six other portfolios. The explanation is a rapid recovery of the BAC stock prices in the post-crisis period. Another reason is a lesser diversification, since TR1 is a risky trading rule, as compared with others, used in this research. AAPL and NOK stocks were also present, but in much smaller numbers. The largest MAE was obtained using the AR(9) forecast model, and the greatest MSE was shown by the AR(6) model. All the AR-ABS models provided almost equal errors similar to the minimal error, obtained by the AR1 model. This fact can be explained by "weaker fitting" of AR-ABS models, minimizing the absolute errors instead of the quadratic ones. However, in this setup, the AR-ABS(6) model, which provides close to the minimal prediction error, was the only model that delivered losses.

A comparison of profits and prediction errors indicates that the minimal prediction errors do not necessarily provide maximal profits. In this case, the maximal profit was achieved by the AR(6) model, the prediction error of which is close to the maximal one. This situation is confirmed by the positive profit-prediction error correlations.

**Table 1** *Average portfolios using ten trading rules in the real stock market*

| | **TR1** | **TR2** | **TR3** | **TR4** | **TR5** | **TR6** | **TR7** | **TR8** | **TR9** | **TR10** |
|---|---|---|---|---|---|---|---|---|---|---|
| | **AR(6)** | **AR-ABS(1)** | **AR(6)** | **AR-ABS(9)** | **AR-ABS(3)** | **AR-ABS(3)** | **AR(6)** | **AR(6)** | **AR(1)** | **AR(6)** |
| **MSFT** | 37,68 | 25,39 | 50,23 | **193,31** | 5,51 | **3,58** | 5,82 | 1,88 | 0,00 | 2,78 |
| **AAPL** | 0,66 | **86,86** | 24,41 | 49,24 | 0,27 | 0,40 | 0,69 | 0,64 | **3,14** | 1,41 |
| **GOOG** | 0,36 | 0,00 | 2,00 | 3,28 | 0,04 | 0,23 | 0,01 | 1,15 | 0,25 | 0,37 |
| **NOK** | 7,24 | 0,05 | 121,13 | 25,42 | 0,23 | 0,46 | 7,38 | 1,96 | 0,00 | 1,78 |
| **TM** | 1,28 | 0,00 | 9,38 | 2,81 | 0,01 | 0,10 | 0,57 | 0,67 | 0,62 | 1,02 |
| **BAC** | **1937,03** | 5,84 | **495,88** | 64,83 | **7,68** | 3,06 | **14,68** | **4,23** | 0,25 | 2,45 |
| **BAC** | 1,14 | 0,01 | 24,96 | 12,33 | 0,03 | 0,16 | 0,25 | 0,43 | 0,17 | 0,86 |
| **ORCL** | 43,43 | 0,09 | 43,41 | 10,77 | 4,10 | 2,64 | 9,22 | 1,08 | 0,00 | **3,22** |
| **Average** | *253,60* | *14,78* | *96,43* | *45,25* | *2,23* | *1,33* | *4,83* | *1,51* | *0,55* | *1,74* |

Figure 2 illustrates the growth tendency of stock prices in the post-crisis economic conditions. An interesting observation is that a similar growth tendency was also noticed in the stock prices, generated by the virtual stock exchange (see Figure 5).

Using the most risky TR1 and different forecast models, traders preferred BAC stocks. AAPL, NOK, and MSFT stocks were also traded. The remaining stocks were traded less or not traded at all. In the Period I, the greatest profit was reached using the prediction model AR(6) with a portfolio mostly of just one stock BAC.

Figure 3 shows how the portfolio changes, using TR4 and the prediction model AR(1).

By comparing Figures 3 and 4, one can see that, using the prediction model AR(9), the trading is irregular and different from the slow trading pattern of model AR(1). The average profit of AR(9) is 4637.71 as compared with 5579.59 of the AR(1) model. A possible explanation of the lesser profit of AR(9) is higher transaction costs due to frequent buying-selling operations. This fact shows that, in the short time investment, smoothness of predictions is an important factor, too.

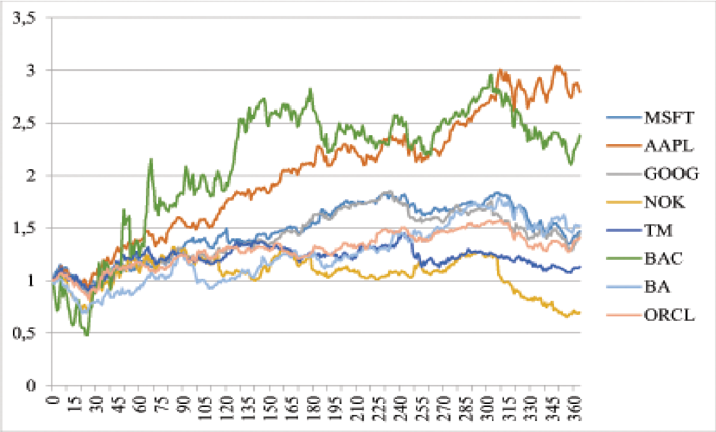


**Figure 3** *Daily prices of eight stocks in the real market, Period I*

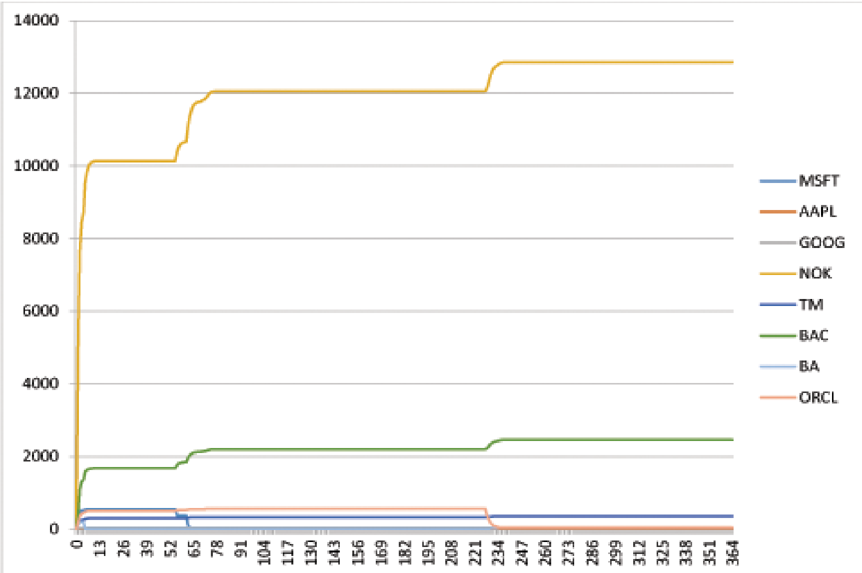


**Figure 4** *Portfolio changes in the real stock market, using TR4 and AR(1), Period I*

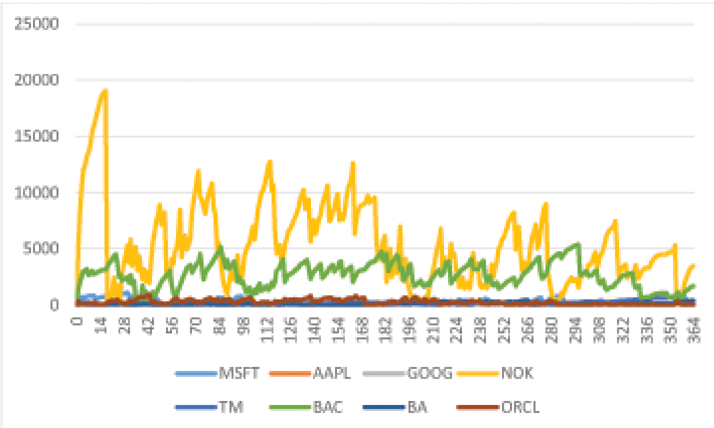


**Figure 5** *Portfolio changes in the real stock market, using TR4 and AR(9), Period I*

### *Virtual stock experiment*

In this section, the experiments, using virtual data and four short-term trading rules, are discussed. Here the greatest profit was obtained by the strategy TR3, AR(9) and the largest losses occurred using TR1, AR-ABS(3).

The largest errors occurred using AR(6) and AR(9), as in the real examples. It shows that, in the virtual environment, simple models provide smaller errors, too.

Table 2 shows the average portfolios of eight prediction modes and four trading rules. Here all stocks are included in all the portfolios, but most popular are the second, sixth, and eighth stocks.

**Table 2** *Average portfolios of four trading rules in the virtual stock market*

| | TR1 | TR2 | TR3 | TR4 |
|---|---|---|---|---|
| | **AR(6)** | **AR(6)** | **AR(9)** | **AR(3)** |
| **first** | 256,00 | 82,00 | 208,00 | 398,00 |
| **second** | **1889,00** | 471,00 | **1518,00** | 663,00 |
| **third** | 23,00 | 73,00 | 57,00 | 58,00 |
| **fourth** | 2,00 | 84,00 | 47,00 | 629,00 |
| **fifth** | 98,00 | 630,00 | 313,00 | 482,00 |
| **sixth** | 565,00 | 48,00 | 91,00 | 9,00 |
| **seventh** | 125,00 | 171,00 | 68,00 | 493,00 |
| **eighth** | **1338,00** | 540,00 | **1326,00** | **1238,00** |
| **Average** | *537,00* | *262,38* | *453,50* | *496,25* |

Fig. 6 shows normalized daily average prices of eight different stocks in the virtual market. Comparing this graph with the historical prices in the post-crisis period, see Fig. 3, some similarity can be noticed.

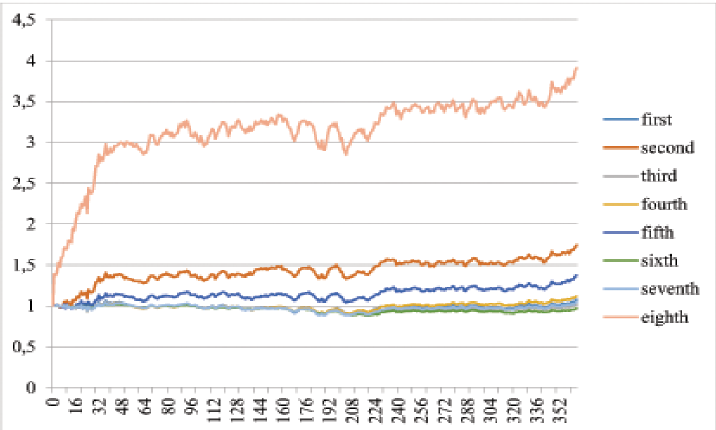


**Figure 6** *Normalized average prices of eight different stocks in the virtual market, using TR1*

### *Real and virtual stock experiment – average results*

To see the general pattern not covered by details, the average results will be presented. The visual comparison does not indicate an obvious positive or negative relation of profits and prediction errors. That visual impression is corroborated in the following section of correlations.

Fig. 7 shows the average profits and prediction errors of ten trading rules in the Period I.

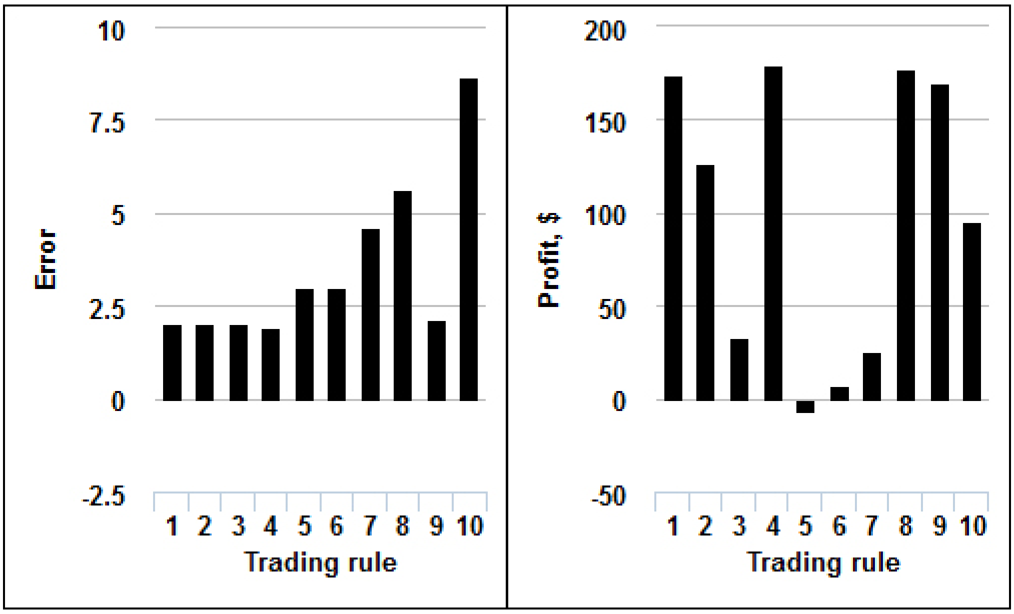


**Figure 7** *The average profits and prediction errors, Period I*

Fig. 8 shows the average profits and prediction errors of four trading rules in the virtual stock exchange.

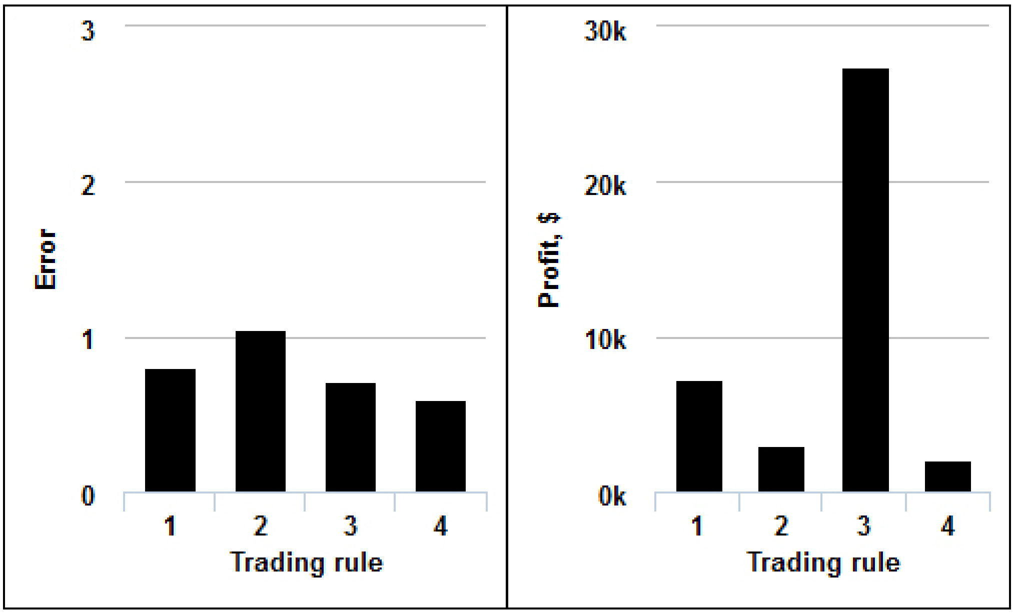


**Figure 8** *The average profits and prediction errors in the virtual stock exchange*

***Real stock experiment – a detailed comparison***

Fig. 9 and 10 are selected to show details on the relation of portfolio profits and prediction errors of asset prices. No general pattern is observed, so a set of diagrams with minimal comments will be shown to render an opportunity to see and evaluate the experimental results directly. In the charts, the profits and prediction errors are shown for different prediction models.

Fig. 9 shows the profits and prediction errors for different prediction models using TR1 in the Period I.

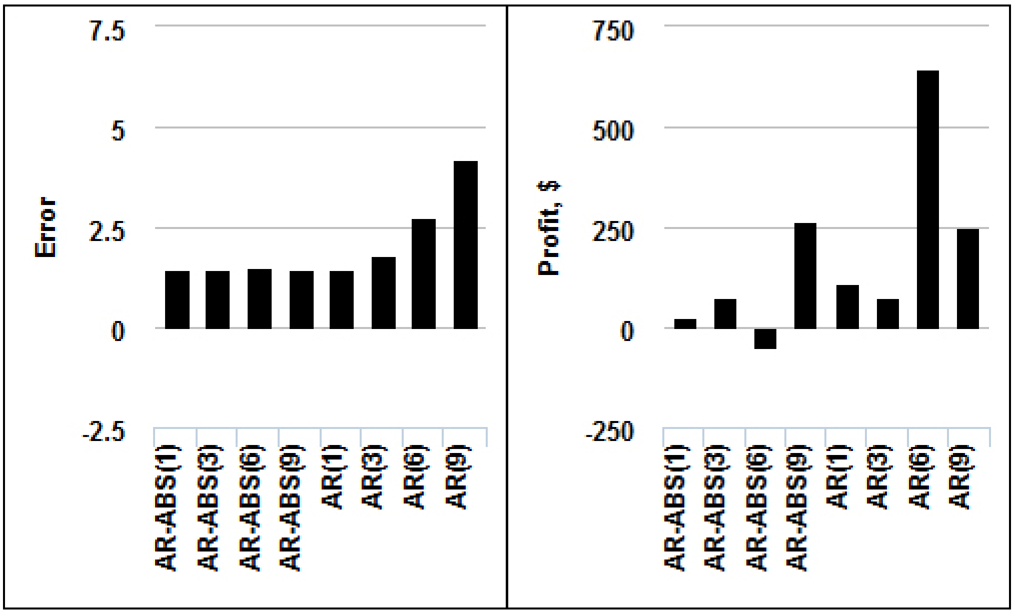


**Figure 9** *Profits and prediction errors for different stocks, using TR1, Period I*

Fig. 10 shows the profits and prediction errors of different prediction models, using TR8 in the Period III.

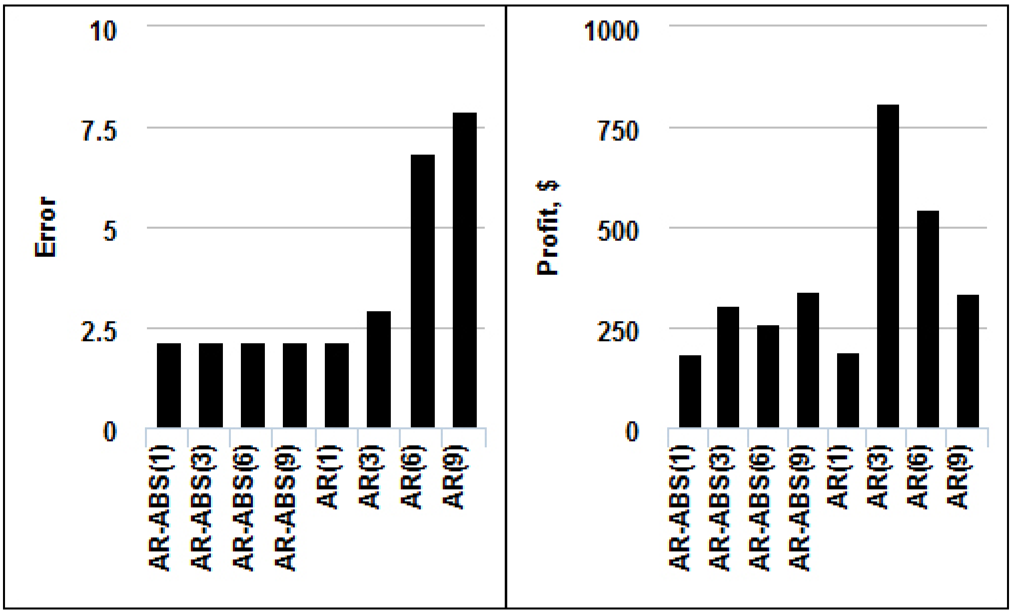


**Figure 10** *Profits and prediction errors for different stocks, using TR8, Period III*

***Virtual stock experiment – a detailed comparison***

To visualize profit and price prediction errors in the virtual environment, a set of charts is presented. In the charts, the profits and prediction errors are shown of different prediction models.

Fig. 11 shows the profits and prediction errors of different prediction models using TR1.

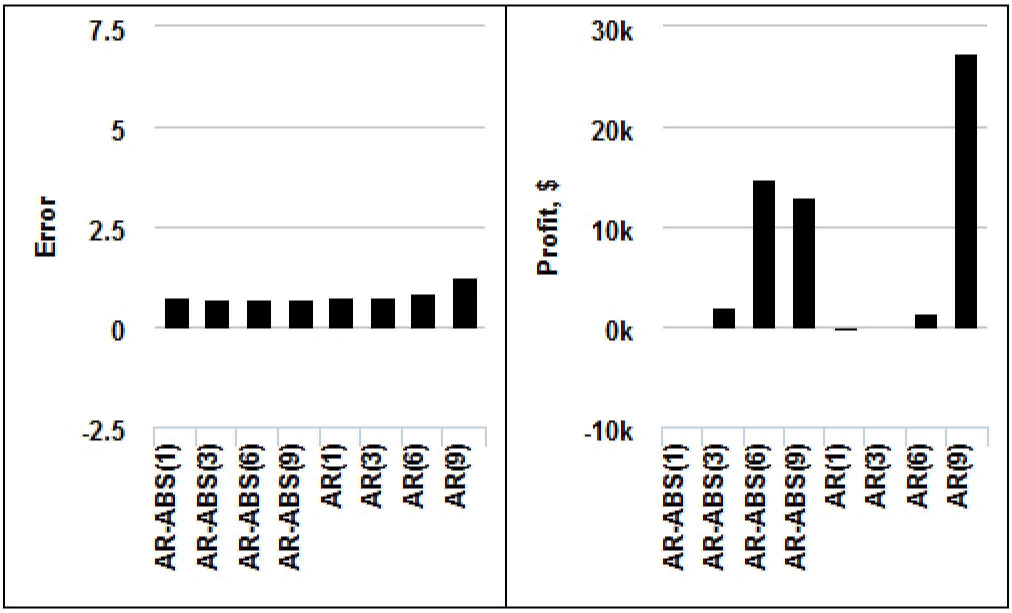


**Figure 11** *Profits and prediction errors for different stocks, using TR1*

Fig. 12 shows the profits and prediction errors for different prediction models using TR4.

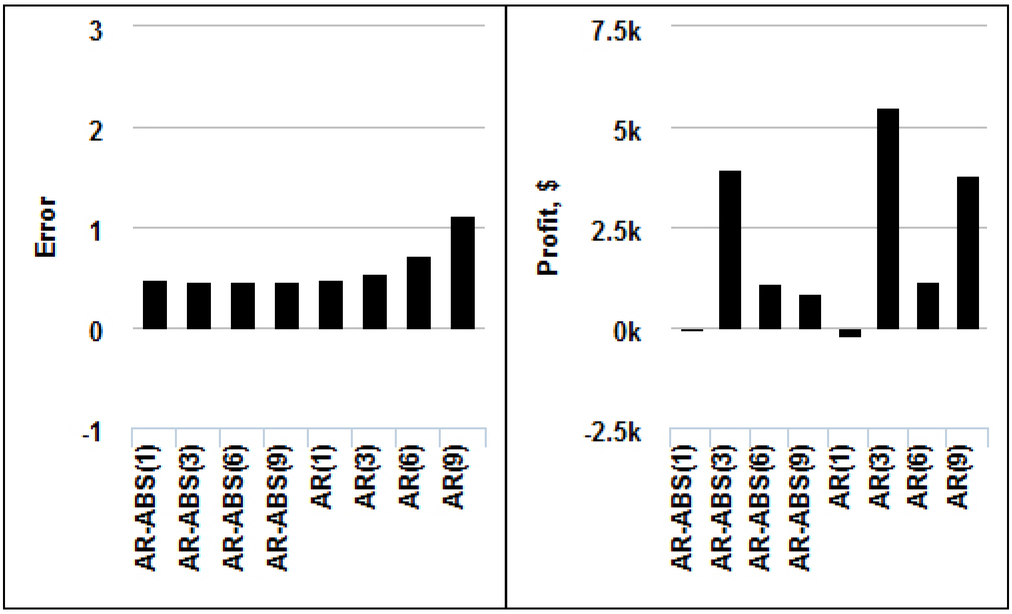


**Figure 12** *Profits and prediction errors for different stocks, using TR4*

The chart and tables helps to investigate the relation of profits and prediction errors by direct visual inspection. The main advantage is that average errors can be compared with the "actual" profits, calculated for entire periods including all the major factors, simulating the real life investment process. A disadvantage is that no numerical value can be attached to measuring the relation strength. This is done in the next section on correlations.

However, the standard correlation formulas are for a set of pairs of related random variables, in this case, for the pairs of daily profits and prediction errors. So, the disadvantage is that the final profit of the whole period is replaced by a sequence of daily profits. It means that not the initial trading rules are simulated, but some their modifications.

***On the correlation between the portfolio profits and price prediction errors, separately for different stocks***

In this work, comparing the figures showing prediction errors with that representing the average profits, it was noticed that minimal MAE and MSE do not necessarily provide the maximal profits. This observation contradicts the general opinion that the investors, who predict stock prices better, are rewarded by higher profits.

To investigate this intuitively paradoxical situation in a formal way, correlations between the profits and prediction errors with the confidence intervals (Shen & Lu, 2006) were calculated. The correlations were determined for each stock separately. They were different, because the prediction errors depend on the asset. The daily correlations cannot be defined for longer term rules, since, by these rules, profits are calculated only once for the whole period. The stocks are obtained at the end of the learning set and sold at the end of the testing set. Therefore, only four short-term trading rules will be investigated.

On the basis of the visual comparison of profits and prediction errors, a small but mainly negative correlation between the profits and standard prediction errors was established. However, only the first part of this assumption was realized. Unexpectedly, the positive correlation happened and, occasionally, the differences were statistically significant at the 95% level of confidence interval.

It is not an easy task to define exact reasons of the observed week and sometimes a reverse relation between the actual profits and the price prediction accuracy. However, the reverse relation is not statistically significant when calculating the averages of all the stocks in the portfolio.

The only clear conclusion is that the direct prediction of profits of a particular trading rule should be the main objective of an investor. The accuracy of predictions is just one of the factors. The following figures (with minimal comments) are to illustrate these statements. None general pattern was observed, so a set of diagrams will be shown to render an opportunity for the readers to see and evaluate the experimental results directly, hoping for some feedback.

In Fig. 13, the negative correlations dominate, mainly MSFT, AAPL, and GOOG. In this way, the Period I, describing the post-crisis growth tendency differs from the more stable Periods II and III. Apparently, in the presence of the well-defined positive trend, the role of prediction accuracy is more important. Thus, the results do not contradict the intuitive beliefs. That is corroborated by the following results of other short-term trading rules in Period I.

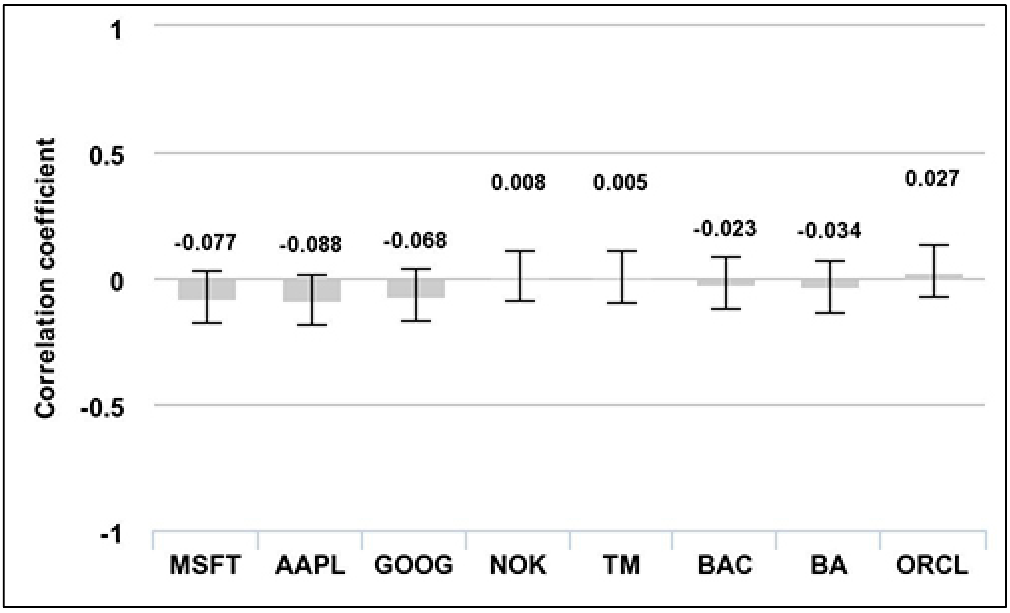

**Figure 13** *Correlations of profits and prediction errors, using TR1, Period I*

In Fig. 14, the majority of correlations are positive, indicating the reverse profit-prediction accuracy relation in the Period II with more stable economic conditions. However, the corresponding average correlation is very small and can be explained by random factors. Comparing with the profit-price change correlations, one observes different results: most correlations are significantly positive.

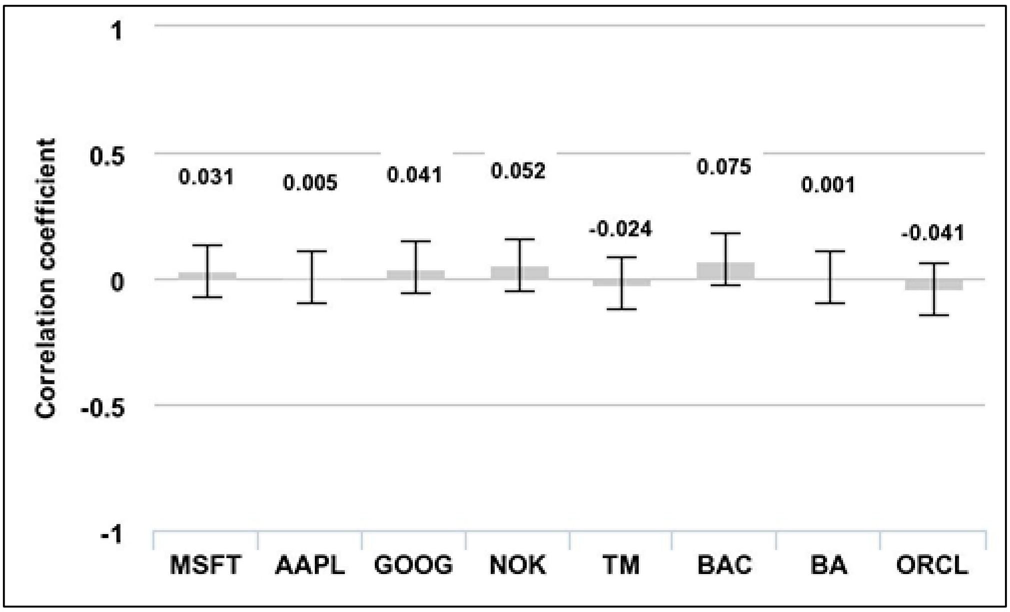

**Figure 14** *Correlations of profits and prediction errors, using TR1, Period II*

***Average correlations between the profits, price prediction errors, and stock price changes***

In this section, the average correlations of different data are investigated with a view to explain the reasons of a weak and occasionally negative relation between the profits and accuracy of price predictions. The average correlations of all the stocks in the portfolio are shown. The average correlations are defined as correlations between the portfolio profits and average price prediction errors. The averages represent the properties of the whole portfolio, not its individual stocks. The average of errors without some correction of their importance means an assumption that the contributions of different stocks are equal. This fact may yield some peculiar results. However, the assumption is needed, at least at the first stage of investigation, to avoid additional complications.

It a difficult to explain directly the weak and positive correlation between the profits and price prediction errors, since the relation of profits and prediction errors is non-linear and involves many additional conditions. The relation between price changes and prediction errors is simpler and easier for understanding, especially using the simplest prediction models such as RW.

However, the relation of price changes and profits has to be defined in order to obtain the final result, – the correlation of profits and prediction errors. In the absence of transactions, the profits are linear functions of stock price changes. In the active market, profits depend on trading rules including additional conditions. Explicit expressions are too complicated, so the profits are determined as respective correlations. However, correlations are intended to define a linear statistical dependency. The trading rules involving additional conditions distort the linearity. Therefore, in some cases, is unexpected results are noticed. However, the linearity assumption helps to grasp some general tendencies better.

In the Periods I and IV, the profit and price change correlations happen to be significantly positive. In the Periods II and III, the correlations are also positive, but not all significantly.

The relation of profits and prediction errors can be determined assuming that, on average, the profits are an increasing function of price changes (differences between the present and past prices). An unexpected result is that significant correlations of price changes and prediction errors are transformed into mostly insignificant correlations of profits and prediction errors.

Thus, in the first step, the correlations of price changes and prediction errors were defined. In the second step, the correlations of price changes and profits were calculated. The profit and prediction error correlations were determined in the final step.

The essential part of the discussion is the explanation of the weak and positive correlations between price changes and prediction errors. The price changes are defined as differences between the present and past prices. The prediction deviations are differences between the predicted and actual present prices. The prediction errors are absolute or squared deviations.

According to the simplest RW prediction model, that represented the well-known theory of efficient market (Fama, 1995), the predicted price is equal to today’s price. It means that the prediction errors are equal to absolute or squared price changes. If the time series can also be described by the RW model, then the positive and negative components in the correlation formula will be approximately equal and the correlation will be almost zero.

In line with the RW model, prediction errors are absolute or squared price changes. Therefore, the positive components of correlations between price changes and price prediction errors prevail. Thus, the positive correlation of price changes and price prediction errors was observed. It produces positive profit-prediction error correlations, as usual, because, on average, the profits are an increasing function of price changes. Such time series could be called as RW-unpredictable, since one cannot predict the prices correctly, using the RW model.

However, using autoregressive models AR($p > 1$) in time series with well-defined trends of price changes, the opposite results are expected. For example, using these models in periods with the positive price trends, the positive price changes produce smaller average prediction errors as compared with the negative changes. The differences would be most obvious when the price volatility is relatively small and AR($p > 1$) models are well-fitted. Thus, the time series that are RW-unpredictable would be predictable using the autoregressive models AR($p > 1$), as usual.

An important question is why the AR-ABS models behave like the RW models and unlike the AR($p > 1$) ones. A possible reason is that, in time series without well-defined price trends, the result $a_1 = 1$, $a_i = 0$, $i = 1, \ldots, p$ of minimizing absolute deviations in AR-ABS by the Simplex algorithm of Linear Programming can be achieved with a higher probability than using AR models when solving the respective systems of linear equations. Another reason is that in AR-ABS, the parameter $p$ is automatically reduced to a unit when at least one of the determinants in the sequence of Gaussian transformations in the Simplex algorithm is too small. On the contrary, using AR, this transformation is performed only once, so the reduction of $p$ is less probable. Note that both models AR-ABS($p > 1$) and AR($p > 1$) can be regarded as an approximation of the AR model.

Statistically significant and negative correlations between the average price changes and prediction errors were observed in time series with well-defined price trends. In the last two periods, the positive correlations prevailed. The Period II shows mixed results: one statistically significant and positive correlation, and two statistically significant and negative correlations. In Period IV, the correlations of average price changes and prediction errors happen to be positive. In contrast, in the Period I, the respective correlations are significantly negative.

Since on average the profits are increasing functions of price changes, the positive, mostly statistically significant respective correlations were observed in all the periods. This fact explains the importance of correlations between the price changes and prediction errors in defining the final result: the correlations of profits and prediction errors. However, there was an unexpected exception: in the Period IV with the positive correlations of price changes and prediction errors, statistically insignificant and negative correlations between profits and prediction errors were noticed.

The explanation is that in the AR($p > 1$) models, the correlations between price changes and prediction errors were also negative. A possible reason why the three negative correlations of AR($p > 1$) models outweigh the five positive ones is that the correlations of profit and price changes of the negative components are larger. Another reason is that, when defining the correlations, the average errors of all stocks are regarded as equally important with no distinction between the active stocks and stocks almost absent in the portfolio. An illustration is the Nokia stocks.

The negativity and positivity of the correlation between price changes and prediction errors don't necessarily determine the negativity and positivity of correlations between profits and prediction errors, since the correlations are defined by the interaction of actual price changes of different stocks and predictions of different models that often produce the opposite results.

A source of uncertainty is that in the multi-stock case, the profit is defined for the whole portfolio, but not for individual stocks, while the prediction errors are calculated separately for each stock. Note that profit depends on two factors: the prediction model and the trading rule. The experiments of this work indicate that the trading rule is the most important factor. Additional explanations will be provided commenting the respective figures.

Fig. 15 shows the correlations of profits and price changes. All the correlations are statistically significantly and positive, and depend on the trading rules, which is quite natural, since profits depend on the trading rules and are increasing functions of price changes, on average. It means that price changes are good predictors of portfolio profits.

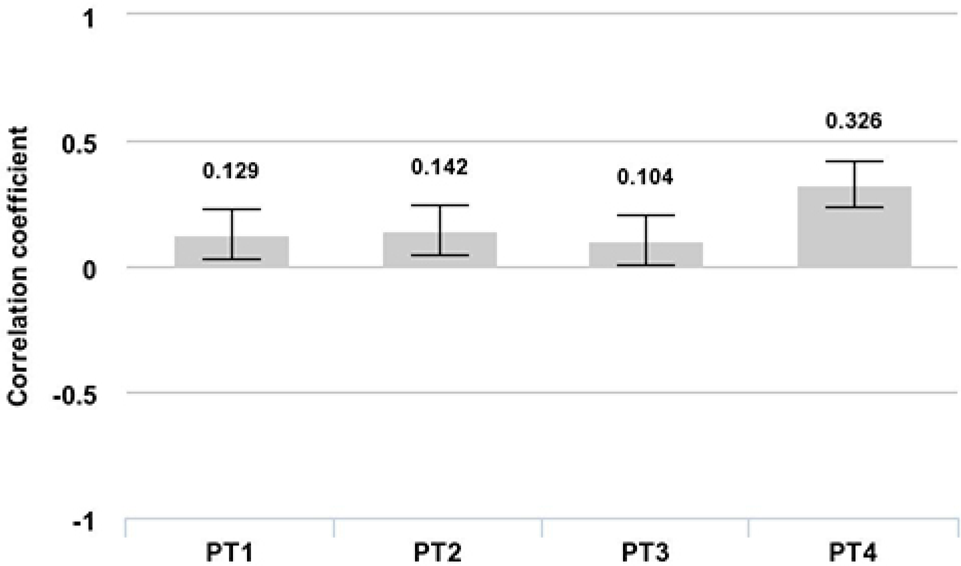


**Figure 15** *Correlations of profits and price changes, using different trading rules, Period I*

Unfortunately, price changes can be predicted reliably only in the periods of well-defined price trends. In the stable economic periods, price changes can be close to RW model. It explains the weak and often positive correlations between profits and prediction errors in the Periods II and III.

Fig. 16 shows the correlations of price changes and prediction errors (averages of different stocks) in the Period I, using TR1. There are two different groups of models: the AR-ABS and AR(1) models are in the first group, the remaining three AR($p > 1$) models belong to the second group. The same grouping remains during all the four periods. However, different groups behave differently depending on time. During the Period I, the first group, correlations were mostly statistically significantly and negative. In the second group, correlations were also negative, but small and statistically insignificant.

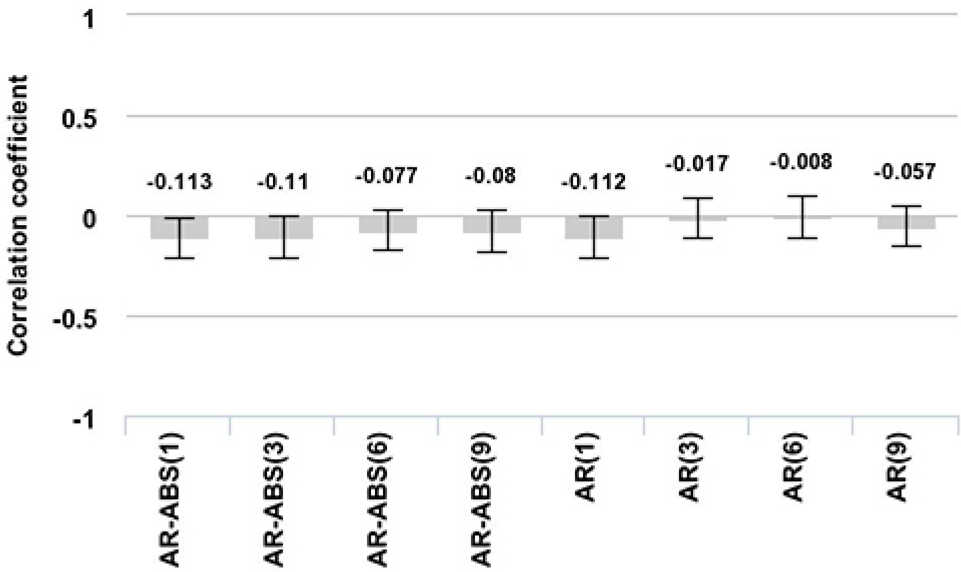


**Figure 16** *Correlations of price changes and prediction errors, using different prediction models and using TR1, Period I*

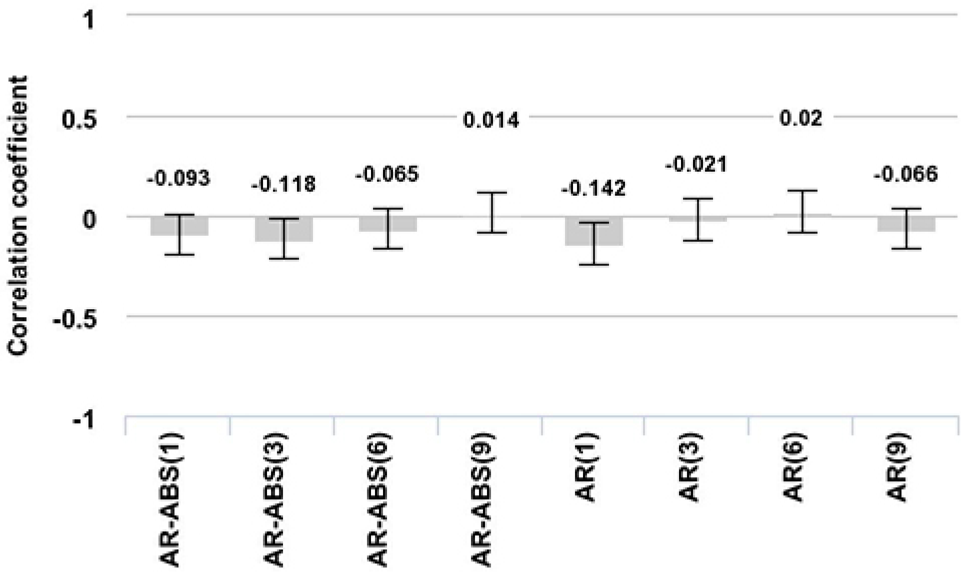


**Figure 17** *Correlations of profits and prediction errors, using different prediction models and using TR3, Period I*

Fig. 17 shows the correlations of profits and prediction errors (averages of different stocks) in the Period I, using TR3. Unexpectedly, the image is different from that of price changes and prediction errors (see Figure 15). A possible explanation is that correlations show linear statistical dependencies, while both the profits and prediction errors are nonlinear functions.

## 5. Conclusions

The research, completed in this thesis, has led to the following conclusions:

1. The main scientific novelty of individual work is application of the stock exchange model (virtual stock exchange) in prediction problem solving and empirical fact explanation that the correlation of portfolio profit and stock prices prediction accuracy is not important statistically and could be negative.
2. The correlation between the portfolio profit and stock price prediction error is statistically insignificant, since the portfolio profit and prediction error are different nonlinear functions. It can be regarded as a special case of over-fitting. The portfolio profit is an increasing (on average) function of price changes (differences between the present and past asset prices) while prediction errors are squared or absolute differences between the predicted and actual asset prices.
3. The correlation of the portfolio profit and price changes is positive and statistically significant, as usual.
4. Predictions of price changes are trust-worthy, if the trend of asset prices is well defined. Predictions of price changes are not reliable without some additional information in the economic periods without the well-defined price trends, because the asset prices can be close to the Random Walk (RW) model.
5. In the short-time investment, not only prediction accuracy, but also its stability is an important factor helping to reduce the transaction costs.
6. An unexpected result is the positive, under some conditions, correlation between price changes and prediction errors using daily data. In the Period IV, the positivity can be disregarded as statistically insignificant. However, in the Periods II and III, the positivity is statistically significant.
7. Positive correlations happen to be using such prediction models, in which greater price changes generate higher errors thus creating positive components of profit and prediction error correlations. Namely that explains occasional positive correlations of portfolio profits and stock price prediction errors, since on average the profit is an increasing function of price changes. A simple example of such a model is Random Walk (RW). Approximate models are AR(1) and AR-ABS(1).
8. A further experimental and theoretical research is desirable to explain the cases of the weak and sometimes negative relation of profits and price prediction accuracy with a view to define reasons and conditions, where the profit is an increasing function of the prediction accuracy and to determine a set of other portfolio profitability factors, in addition to the prediction accuracy.

9. Most of the results of this work have been obtained with the help of the expert system (Virtual Stock Exchange). To do all that by direct online experimentation would be hardly realizable, since the past financial time series cannot be repeated in the real life.

## The list of literature referenced in this summary


Ait-Sahalia, Y. & Hansen, L. P., 2009. *Handbook of Financial Econometrics.* Netherlands: North Holland.

Akerlof, G. A. & Shiller, R. J., 2009. *Animal Spirits:How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism.* Princeton: Princeton University Press.

Allen, F., 2008. *Market illiquidity and financial instability,* s.l.: Federal Reserve Academic Consultants Meeting on Friday, October 3, 2008 on "Financial Stability and the Linkages between the Financial and Real Sectors".

Anderson, E. W., Hansen, L. P. & Sargent, T., 2012. Small noise methods for risk-sensitive/robust economies. *Journal of Economic Dynamics and Control,* 36(4), pp. 468-500 .

Arellano, M., Hansen, L. P. & Sentana, E., 2012. Underidentification?. *Journal of Econometrics,* 170(2), p. 256–280.

Bagočius, V., Zavadskas, E. K. & Turskis, Z., 2014. Multi-person selection of the best wind turbine based on the multi-criteria integrated additive-multiplicative utility function. *Journal of Civil Engineering and Management,* 20(4), pp. 590-599.

Bailey, D. H., Borwein, J. M., de Prado, M. L. & Zhu, Q. J., 2014. Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance. *Notices of the American Mathematical Society,* pp. 458-471.

Bernanke, B. S., 2004. *Essays on the Great Depression.* Princeton, New Jersey: Princeton University Press.

Black, F. & Scholes, M., 1973. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy,* 81(3), p. 637–654.

Borovička, J., Hansen, L. P., Hendricks, M. & Scheinkman, J. A., 2011. Risk Price Dynamics. *Journal of Financial Econometrics,* 9(1), pp. 3-65.

Cochrane, J. H., 1997. *Time Series for Macroeconomics and Finance.* Chicago: Graduate School of Business, University of Chicago.

Dadelo, S., Turskis, Z., Zavadskas, E. K. & Dadelienė, R., 2014. Multi-criteria assessment and ranking system of sport team formation based on objective-measured values of criteria set. *Expert Systems with Applications,* 41(14), p. 6106–6113.

Darley, V. & Outkin, A. V., 2007. *Nasdaq Market Simulation: Insights on a Major Market from the Science of Complex Adaptive Systems.* Hackensack, New Jersey: World Scientific Publishing Company.

Fama, E. F., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance,* 25(2), pp. 383-417.

Fama, E. F., 1995. Random Walks in Stock Market Prices. *Financial Analysts Journal,* pp. 75-80.

Fishburn, P. C., 1964. *Decision and Value Theory.* New York: John Wiley & Sons, Inc..

Greenspan, A., 2009. *Market crisis 'will happen again'.* [Online] Available at: http://news.bbc.co.uk/2/hi/business/8244600.stm [Accessed 25 10 2013].

Hansen, L. P., 2012. Dynamic Valuation Decomposition Within Stochastic Economies. *Econometrica,* 80(3), pp. 911-967.

Hansen, L. P., 2012. Proofs for large sample properties of generalized method of moments estimators. *Journal of Econometrics,* 170(2), p. 325–330.

Hansen, L. P., 2013. Risk Pricing over Alternative Investment Horizons. *Handbook of the Economics of Finance,* 2(B), p. 1571–1611.

Hansen, L. P. & Sargent, T. J., 2007. *Robustness.* Princeton: Princeton University Press.

Hansen, L. P. & Sargent, T. J., 2011. Robustness and ambiguity in continuous time. *Journal of Economic Theory,* 146(3), p. 1195–1223.

Hansen, L. P. & Scheinkman, J. A., 2012. Pricing growth-rate risk. *Finance Stoch,* 16(1), pp. 1-15.

Hansen, L. P. & Scheinkman, J. A., 2012. Recursive utility in a Markov environment with stochastic growth. *Proceedings of the National Academy of Sciences of the United States of America,* 109(30), p. 11967–11972.

Katin, I., 2014. *On Development and Investigation of Stock-Exchange Model.* Vilnius: Vilnius University.

Krugman, P., 1996. *The Self Organizing Economy.* Oxford: Wiley-Blackwell.

Krugman, P., 2000. *Currency Crises (National Bureau of Economic Research Conference Report).* Chicago: University of Chicago Press.

Krugman, P., 2008. *The Return of Depression Economics and the Crisis of 2008.* New York: W. W. Norton & Company.

Krugman, P., 2009. The Increasing Returns Revolution in Trade and Geography. *American Economic Review,* 99(3), p. 561–571.

Krugman, P., 2012. *End This Depression Now!.* United States: W. W. Norton & Company.

Markowitz, H. M., 1952. Portfolio Selection. *The Journal of Finance,* 7(1), pp. 77-91.

Markowitz, H. M., 1959. *Portfolio Selection: Efficient Diversification of Investments.* New York: John Wiley & Sons, Inc..

Merton, R. C., 1971. *Theory od Rational Option Pricing.* Cambridge: Massachusetts Institue of Technology.

Merton, R. C., 1972. An Analytic Derivation of the Efficient Portfolio Frontier. *The Journal of Financial and Quantitative Analysis,* 7(4), pp. 1851-1872.

Mockus, J., 2003. Stock exchange game model as an example for graduate level distance studies. *Computer Applications in Engineering Education,* pp. 229-237.

Mockus, J., 2006. Investigation of Examples of E-Education Environment for Scientific Collaboration and Distance Graduate Studies, Part 1. *Informatica,* 17(2), pp. 259-278.

Mockus, J., 2010. On Simulation of Optimal Strategies and Nash Equilibrium in the Financial Market Context. *Journal of Global Optimization,* pp. 129-143.

Mockus, J., 2012. On Simulation of the Nash Equilibrium in the Stock Exchange Contest. *Informatica,* pp. 77-104.

Mockus, J., 2014. *Global Optimization.* [Online] Available at: http : //optimum2.mii.lt [Accessed 18 01 2014].

Mockus, J. et al., 1997. *Bayesian Heuristic Approach to Discrete and Global Optimization.* Dordrecht-London-Boston: Kluwer Academic Publishers.

Mockus, J., Katin, I. & Katina, J., 2014. On the Optimization of Investment Strategies in the Context of Virtual Financial Market by the Individual Approach to Risk. 25(2).

Mockus, J. & Raudys, A., 2010. On the Efficient-Market Hypothesis and stock exchange game model. *Expert Systems With Applications,* pp. 5673-5681.

Nash, J., 1951. Non-Cooperative Games. *The Annals of Mathematics,* 54(2), pp. 286-295.

Scholes, M. S., 1998. Derivatives in a Dynamic Environment. *American Economic Review,* 88(3), pp. 350-370.

Sharpe, W. F., 1963. A Simplified Model for Portfolio Analysis. *Management Science,* 9(2), pp. 277-293.

Sharpe, W. F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance,* 19(3), pp. 425-442.

Sharpe, W. F., 1966. Mutual Fund Performance. *The Journal of Business,* 39(1), p. 119–138.

Sharpe, W. F., 1994. The Sharpe Ratio. *Yhe Journal of Portfolio Management,* 21(1), pp. 49-58.

Sharpe, W. F., 2000. *Portfolio Theory and Capital Markets.* s.l.:McGraw-Hill Trade.

Sharpe, W. F., 2007. Expected Utility Asset Allocation. *Financial Analysts Journal,* 63(5), pp. 18-30.

Shen, D. & Lu, Z., 2006. *Computation of Correlation Coefficient and Its Confidence Interval in SAS.* San Francisco, California, SAS Institute Inc..

Shiller, R. J., 1989. *Market Volatility.* Cambridge: The MIT Press.

Shiller, R. J., 1998. *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks.* Oxford: Oxford University Press.

Shiller, R. J., 2007. Low Interest Rates and High Asset Prices: An Interpretation in Terms of Changing Popular Economic Models. *Brookings Papers on Economic Activity,* 38(2), pp. 111-134.

Shiller, R. J., 2008. *The Subprime Solution: How Today's Global Financial Crisis Happened, and What to Do about It.* Princeton: Princeton University Press,.

Shiller, R. J., 2009. *Irrational Exuberance.* Princeton: Princeton University Press.

## List of publications on topic of dissertation

### *The articles published in the peer-reviewed journals*

1. J. Mockus, I. Katin, J. Katina. On the experimental investigation of investment strategies in the real and virtual financial markets. Informacijos mokslai / Vilniaus universitetas. 2013, t. 65. ISSN 1392-0561 p. 103-110.
2. J. Mockus, I. Katin, J. Katina. On the Optimization of Investment Strategies in the Context of Virtual Financial Market by the Individual Approach to Risk. Informatica, 2014, vol. 25, issue 2, ISSN 0868-4052.

### *The articles in other scientific journals*

1. J. Mockus, J. Katina, I. Katin. On autoregressive moving-average models as a tool of virtual stock-exchange: experimental investigation. Lietuvos matematikos rinkinys. LMD darbai. 2012, t. 53, ser. A. ISSN 0132-2818 p. 129-134.
2. J. Mockus, I. Katin, J. Katina. On experimental investigation of the web-based stock-exchange model. Lietuvos matematikos rinkinys. LMD darbai. 2012, t. 53, ser. A. ISSN 0132-2818 p. 123-128.

## About the author

Joana Katina was born on the 28th of May, 1982 in Vilnius, Lithuania. In 2000, she graduated from the Vilnius Mykolas Biržiška Gymnasium. She received her bachelor’s degree in Informatics after graduating from the Faculty of Fundamental Sciences at the Vilnius Gediminas Technical University in 2006. She received a master’s degree in Informatics Engineering at the same faculty in 2009. From 2009 till 2014 she studied technological science for a doctoral degree at the Institute of Mathematics and Informatics of Vilnius University.

## PROGNOZAVIMO PROBLEMŲ TYRIMAS VIRTUALIOJE AKCIJŲ BIRŽOJE

### *Tyrimų sritis ir problemos aktualumas*

Atlikta daug mokslinių tyrimų su įvairiais prognozavimo modeliais ir metodais: AR($p$) – autoregresinis, ARMA($p,q$) – autoregresinis slenkančio vidurkio, ARIMA($p,d,q$) – autoregresinis integruoto slenkančio vidurkio, ARFIMA($p,d,q$) – autoregresinis trupmeniškai integruoto slenkančio vidurkio, ARCH($q$) – autoregresinis sąlyginio heteroskedastiškumo, ANN – dirbtinio neuroninio tinklo ir kt. Sudėtingesni prognozavimo modeliai įprastai geriau prisitaiko prie istorinių duomenų (duoda mažesnes paklaidas ankstesnių duomenų atžvilgiu). Vis dėlto ekonomikos sąlygos keičiasi, tad kyla klausimas, kaip šie modeliai prognozuos būsimus finansinius duomenis.

Dažnai manoma, kad geras ankstesnių duomenų pritaikymas finansiniams modeliams, nors ir negarantuoja tikslios būsimų duomenų prognozės, tačiau leidžia tikėtis bent jau teigiamos koreliacijos. Kad ir kaip būtų, naujausi duomenys rodo, kad net esant palankioms sąlygoms tik po labai ilgų laiko eilučių gaunama teigiama koreliacija. Laiko reikia dar daugiau, kai taikomi sudėtingesni modeliai. Šis poveikis aprašomas bendru terminu „persimokymas“, perimtu iš ANN srities.

Kainų prognozavimo metodai yra pagrįsti prielaida, kad dėl tikslesnių kainų prognozių galima gauti didesnį pelną. Matematiškai tai reiškia, kad pelnas yra didėjanti akcijų kainų prognozavimo tikslumo funkcija. Tačiau ši prielaida nebūtinai yra teisinga, ypač sprendžiant daug akcijų apimantį portfelio optimizavimo uždavinį, kai esant įvairioms sąlygoms naudojamos sudėtingos investavimo strategijos. Investavimo strategija apibrėžiama kaip prekybos taisyklės (pirkimo ir pardavimo sąlygos) ir prognozavimo modelio pora. Tam tikromis sąlygomis portfelio pelnų ir prognozuojamų kainų paklaidų koreliacija gali būti teigiama. Tokia koreliacija reiškia, kad portfelio pelnų priklausomybė nuo prognozuojamų akcijų kainų tikslumo yra neigiama.

Šis teiginys yra iliustruojamas keliais eksperimentais. Tai natūralus tokių prognozavimo modelių, kuriuose didesni skirtumai tarp dabartinių ir būsimų akcijų kainų duoda didesnius prognozuojamų kainų nuokrypius, rezultatas. Šiuo būdu galima gauti tiek teigiamą, tiek neigiamą koreliaciją tarp kainų pokyčių ir prognozuojamų kainų paklaidų. Pavyzdžiui, jeigu vyrauja teigiami pokyčiai, gaunama teigiama koreliacija tarp portfelio pelnų ir prognozuojamų kainų paklaidų.

Kitas šio nenuspėjamo reiškinio paaiškinimas gali būti tas, kad daugelio akcijų atveju pelnas yra apibrėžiamas visam portfeliui, ne individualioms akcijoms, o prognozavimo paklaida yra skaičiuojama atskirai kiekvienai akcijai. Pelnas priklauso nuo dviejų veiksnių: prognozavimo modelio ir prekybos taisyklės. Šiame darbe aprašyti eksperimentai parodė, kad prekybos taisyklė yra svarbesnis veiksnys.

Tiriant optimalią pelno ir prognozavimo tikslumo priklausomybę susiduriama su dar vienu sunkumu – pelno optimizavimas atliktinas atsižvelgiant į investavimo strategijas. Todėl maksimizuojama funkcija, kuri didėja priklausomai nuo skirtumo tarp esamų ir buvusių kainų didėjimo ir nuo individualių investuotojų sąlygas išreiškiančių apribojimų. Įprasti prognozavimo modeliai neretai minimizuoja prognozuojamų kainų kvadratinius ar absoliutinius nuokrypius nuo esamų kainų. Taikomi skirtingi optimizavimo kriterijai, todėl minimali prognozavimo paklaida nebūtinai reiškia maksimalų pelną.

Mokslinių darbų, aprašančių ir teoriškai paaiškinančių šį akivaizdžiai naują rezultatą, nebuvo rasta, todėl priklausomybę tarp portfelio pelno ir prognozavimo paklaidos apibrėžiančių funkcijų tyrimas yra svarbus tolesniems teoriniams ir eksperimentiniams tyrimams. Pagrindinis tikslas – apibrėžti tas sąlygas, kuriose portfelio pelnas yra monotoninė prognozavimo tikslumo funkcija. Šis darbas atliktas siekiant ištirti tiek persimokymo įtaką, tiek pelnų ir prognozavimo paklaidų priklausomybę. Pastaroji problema yra daug svarbesnė, nes ji vis dar yra nenuodugniai išnagrinėta finansinio optimizavimo sritis.

Siekiant sumažinti aplinkos ir laikinų veiksnių įtaką reikalingas akcijų biržos modelis, kuris imituotų pagrindinių investuotojų sąveiką, kai naudojamos skirtingos investavimo strategijos ir siekiama maksimaliai padidinti pelną. Kad sumažėtų didelis smulkių investuotojų skaičius, jis pateikiamas kaip Gauso triukšmas. Norint objektyviai stebėti persimokymą tiriami keturi skirtingų ekonomikos sąlygų periodai. Įvertinti pelno ir prognozavimo paklaidos priklausomybei apskaičiuotos koreliacijos naudojant tiek realius, tiek virtualius duomenis. Paminėtina, kad šiame darbe koreliacijos apskaičiuotos tarp pelno ir prognozavimo paklaidos, o ne prognozavimo tikslumo, todėl teigiama koreliacija reiškia neigiamą pelno priklausomybę nuo prognozavimo tikslumo.

Geriausios strategijos skiriasi skirtingais laiko periodais ir realiomis bei virtualiomis rinkomis. Paprastesniais prognozavimo modeliais įprastai gaunamos mažesnės prognozavimo paklaidos. Tai atrodo normalu persimokymo kontekste ir atitinka efektyvios rinkos hipotezę.

Naujas ir svarbiausias šio darbo rezultatas yra faktas, kad minimalios prognozavimo paklaidos nebūtinai duoda maksimalų pelną. Teigiama pelno ir prognozavimo paklaidos koreliacija tam tikromis sąlygomis gauta tiek realioje, tiek virtualioje rinkoje. Svarbi tyrimo dalis priklauso nuo virtualios akcijų biržos, nes šis modelis atspindi tik tam tikras pagrindines lošimo taisykles, kurios gali būti kontroliuojamos mokslininkų ir nepriklauso nuo kintančios aplinkos.

### *Tyrimo objektas*

Tyrimo objektas yra imitacinis akcijų biržos modelis (virtuali akcijų birža) ir jo taikymas statistinei priklausomybei tarp prognozavimo paklaidos ir faktinio pelno nustatyti, kai taikomos įvairios akcijų prekybos taisyklės realiose ir virtualiose finansų rinkose.

### *Darbo tikslas ir uždaviniai*

Darbo tikslas – patobulinti sukurtą modelį, imituojantį akcijų biržos darbo procesą ir pritaikyti jį prognozavimo uždavinių tyrimams. Taip pat siekiama taikant šį modelį ištirti priklausomybę tarp pelno ir akcijų kainų prognozavimo tikslumo, tarp realių duomenų ir teorinių modelių.

### *Tyrimo metodai*

Analizuojant mokslinius ir eksperimentinius pasiekimus sprendžiant optimalaus finansinio investavimo (portfelio) uždavinį, įskaitant prognozavimo ir rinkos modelių pasiekimus, naudoti informacijos paieškos, sisteminimo, analizės, lyginamosios analizės ir apibendrinimo metodai. Tiriant priklausomybę tarp portfelio pelnų ir prognozavimo paklaidų naudotas akcijų biržos modelis. Kitos dienos kaina nustatoma taikant autoregresinius prognozavimo metodus. Remiantis eksperimentinio tyrimo metodu, atlikta statistinė duomenų ir tyrimų rezultatų analizė, o jos rezultatams apibendrinti taikytas lyginamosios analizės metodas.

### *Darbo mokslinis naujumas*

Pagrindinis darbo naujumas yra akcijų biržos modelio (virtualios akcijų biržos) pritaikymas prognozavimo uždavinių tyrimams. Taip pat paaiškintas naujas empirinis faktas, kad portfelio pelno ir prognozavimo tikslumo koreliacija yra statistiškai nereikšminga ir gali būti neigiama. Tai reiškia, kad vien tik prognozavimo tikslumas neužtikrina didesnio pelno ir tam tikromis sąlygomis investuotojai, taikantys ne tokius tikslius prognozavimo metodus, gali tikėtis didesnio pelno.

### *Darbo rezultatų praktinė reikšmė*

Išvada, kad vien tik prognozavimo tikslumas neužtikrina didesnio pelno, turi įtakos investavimo optimizavimo tyrimui. Ypatingas dėmesys turėtų būti skiriamas patikimų prekybos taisyklių, nejautresnių nenuspėjamiems rinkos pokyčiams, paieškai. Tobulesnis virtualios akcijų biržos variantas PORTFOLIO2 gali būti taikomas ir kitiems akcijų biržos tyrinėjimams.

### *Ginamieji teiginiai*

Taikant objektyviai egzistuojančias akcijų portfelio formavimo taisykles, koreliacija tarp prognozavimo tikslumo ir faktinio pelno yra silpna ir daugeliu atvejų neigiama tiek realioje, tiek virtualioje finansų rinkose. Tai rodo, kad faktinis pelnas taikant realias prekybos taisykles nebūtinai yra monotoninė statistinė prognozavimo paklaidos funkcija.

### *Darbo rezultatų aprobavimas*

Tyrimų rezultatai publikuoti 5 moksliniuose leidiniuose: 2 recenzuojamuose periodiniuose mokslo leidiniuose, 2 recenzuojamuose kituose mokslo leidiniuose, 1 recenzuojamame konferencijos darbų leidinyje. Tyrimų rezultatai pristatyti ir aptarti 8 nacionalinėse ir tarptautinėse konferencijose Lietuvoje

### *Disertacijos struktūra*

Disertaciją sudaro 5 skyriai, bibliografijos sąrašas ir priedas. Disertacijos skyriai: Įvadas, Tyrimų apžvalga, PORTFOLIO modelis, Eksperimentiniai tyrimai, Bendrosios išvados. Be to, disertacijoje pateikti lentelių, paveikslų, naudotų žymėjimų ir santrumpų sąrašai. Disertacijos apimtis – 115 puslapis (be priedų), kuriuose pateikti 86 paveikslai, 58 formulės ir 3 lentelės. Disertacijoje remtasi 60 bibliografinių šaltinių

*Išvados*

Eksperimentinių tyrimų rezultatai leido padaryti šias išvadas:

1. Darbo pagrindinis naujumas – galimybė modelį taikyti prognozavimo uždavinių tyrimams ir paaiškintas empirinis faktas, kad portfelio pelno ir akcijų kainų prognozavimo tikslumo koreliacija yra statistiškai nereikšminga ir gali būti neigiama.
2. Koreliacija tarp portfelio pelno ir akcijų kainos prognozavimo paklaidos, kai naudojami dieniniai duomenys, yra statistiškai nereikšminga todėl, kad pelnas ir prognozavimo paklaida yra skirtingos netiesinės funkcijos. Portfelio pelnas yra didėjanti (vidurkio prasme) kainų pokyčio funkcija, o prognozavimo paklaidos yra kvadratiniai arba absoliutiniai skirtumai tarp dabartinių ir prognozuojamų kainų.
3. Portfelio pelno ir kainų pokyčių koreliacija paprastai yra teigiama ir statistiškai reikšminga.
4. Kainų pokyčių prognozavimas yra patikimas, jeigu akcijų kainų kitimo tendencija yra užtenkamai išreikšta. Kitu atveju, jei nėra papildomos informacijos, kainų pokyčių prognozavimas nėra patikimas, nes akcijų kainos kinta pagal efektyvios rinkos dėsnius, t. y. panašiai kaip ir pagal RW modelį.
5. Trumpalaikiam investavimui svarbus ne tik prognozavimo tikslumas, bet ir jo stabilumas, nes tai mažina visą prekybos operacijų kainą.
6. Nelauktas naujas rezultatas yra tam tikromis sąlygomis gauta teigiama koreliacija tarp pelno ir kainų prognozavimo tikslumo, kai naudojami dieniniai duomenys. IV periodu teigiama koreliacija nėra statistiškai reikšminga, tačiau II ir III periodais ši koreliacija statistiškai reikšminga.
7. Koreliacijos teigiamos, kai taikomi tokie prognozavimo modeliai, kuriais didesni akcijų kainų pokyčiai generuoja didesnes paklaidas ir taip sukuria teigiamas pelno ir kainų prognozavimo paklaidų koreliacijos dedamąsias. Būtent šis faktas paaiškina teigiamas pelno ir kainų prognozavimo paklaidų koreliacijas, nes pelnas yra vidutiniškai didėjanti kainų prieaugio funkcija. Paprasčiausias tokio modelio pavyzdys yra RW modelis, o į jį panašūs yra AR(1) ir AR-ABS(1).
8. Reikėtų toliau atlikti teorinius ir eksperimentinius tyrimus, kad būtų patikslintos priežastys ir sąlygos, kada pelnas yra didėjanti prognozavimo tikslumo funkcija, ir kad būtų nustatyti kiti, be prognozavimo tikslumo, įtakos pelnui turintys veiksniai.
9. Dauguma šio darbo rezultatų gauti taikant akcijų biržos modelį (virtualią akcijų biržą), atsižvelgiant į tai, kad statistiniai pelno ir paklaidų tyrimai realioje biržoje sunkiai realizuojami – dirbant realiu laiku akcijų kainos nesikartoja.

JOANA KATINA

INVESTIGATION OF PREDICTION PROBLEMS
BY THE VIRTUAL STOCK EXCHANGHE

Summary of Doctoral Dissertation

Technological Sciences,
Informatics Engineering (07 T)

Editor Janina Kazlauskaitė

JOANA KATINA

PROGNOZAVIMO PROBLEMŲ TYRIMAS
VIRTUALIOJE AKCIJŲ BIRŽOJE

Daktaro disertacijos santrauka

Technologijos mokslai,
Informatikos inžinerija (07 T)

Redaktorė Jorūnė Rimeisytė