VILNIUS UNIVERSITY

Igor Katin

ON DEVELOPMENT AND INVESTIGATION OF STOCK-EXCHANGE MODEL

Doctoral Dissertation

Technological Sciences, Informatics Engineering (07T)

Vilnius, 2014

Doctoral dissertation was prepared at the Institute of Mathematics and Informatics of Vilnius University in 2009–2013.

Scientific Supervisor

Prof. Dr. Habil. Jonas MOCKUS (Institute of Mathematics and Informatics of Vilnius University, Technological Sciences, Informatics Engineering – 07T).

VILNIAUS UNIVERSITETAS

Igor Katin

AKCIJŲ BIRŽOS MODELIO SUDARYMAS IR TYRIMAS

Daktaro disertacija

Technologijos mokslai, informatikos inžinerija (07T)

Vilnius, 2014

Disertacija rengta 2009–2013 metais Vilniaus universiteto Matematikos ir informatikos institute.

Mokslinis vadovas

Prof. habil. dr. Jonas MOCKUS (Vilniaus universiteto Matematikos ir informatikos institutas, technologijos mokslai, informatikos inžinerija – 07T).

Acknowledgments

I would like to express my thanks to all the people who have been in one way or another involved in the preparation of this thesis.

First of all, I would like to express my sincere gratitude to my scientific supervisor Prof. Dr. Habil. Jonas Mockus for the continuous support of my research, for his patience, motivation, enthusiasm, and immense knowledge. I could not have imagined having better mentor for my doctoral study.

I want to express my big thanks to Vilnius University Institute of Mathematics and Informatics director Prof. Dr. Habil. Gintautas Dzemyda for granted conditions for doctoral studies. Special thanks to my thesis reviewers Assoc. Prof. Dr. Olga Kurasova and Dr. Aistis Raudys who read and provided constructive feedback regarding this dissertation. I would also like to thank the Lithuanian State Studies Foundation and Lithuanian Academy of Sciences for the financial support for doctoral studies.

Finally, I wish to thank all my relatives, friends and most of all for my wife Joana for all their support and patience during this challenging period of my life.

Igor Katin

Abstract

A simple Stock Market Game Model (SEGM) was introduced in (Mockus, 2002) to simulate the behavior of several stockholders trading a single stock. In (Mockus, 2010; Mockus and Raudys, 2010), the model was investigated and compared with real data.

In contrast, the proposed model PORTFOLIO is simulating stock exchange including a number of different stocks. The objective of PORTFOLIO is not forecasting, but simulation of stock exchange processes that are affected by predictions of the participants. The main improvements are the multi-stock extension and a number of different trading rules, which represent both the heuristics of potential investors and the well-known theoretical investment strategies.

This 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 comparing with traditional single-stock models with 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 (Krugman, 2000, 2008, 2009).

The model has been compared with actual financial time series and found the results to be close in some cases. The model is designed as a tool to represent behavior of individual investor, which 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 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. The unexpected result was that the minimal stock price prediction errors do not necessarily provide the maximal profits. Therefore, the complete information is presented for the independent testing and verification of this important new result.

The experiments with both the historical and virtual time series show that the profitability of investments depends mainly on trading rules, so the optimization should be performed on the set of trading rules by the direct simulation of these rules using the corresponding stock-market models. This partly explains the weak correlation of profits and prediction accuracy.

List of Figures

List of Tables

Notations

Symbols

a	*The coefficient of market inertia*
$AR(p)$	*The autoregressive model of order p*
$b(t,i)$	*The funds borrowed at moment t*
$B(t,i)$	*The borrowed sum of the stockholder i accumulated at time t*
$C_0(t,i)$	*The investors own money at time t*
$D(t,i)$	*The income from selling and buying stocks at time t*
E	*The expected value of the excess of the asset return*
$I(t)$	*The funds invested at time t*
l	*The profitability level*
$L(t,i)$	*The credit limit at time t*
M	*The number of discrete values of wealth*
$n = n(t)$	*The number of transaction stocks*
$p(t,i)$	*The relative profit at time t by the player i*
q_i	*The insolvency probability*
R	*The return on the portfolio*
S	*The Sharpe ratio*
$S(t,i)$	*The buying-selling strategy of the player i at time t*

$u(y)$	*The utility the wealth y*
$U(t) = U(t,i)$	*The actual profit accumulated at time t by the player i*
$U(x)$	*The expected utility function*
$v(i)$	*The standard deviation*
w_i	*The weighting of component asset i*
x	*The capital distribution*
y^k	*The wealth*
$z(t) = z(t,i)$	*The stock price at time t, predicted by the player i*
$Z(t)$	*The actual stock price at time t*
$\alpha(t)$	*The yield at time t*
$\beta(s,i)$	*The accuracy of prediction*
$\beta(t,i)$	*The relative stock price change at time t as predicted by the player i*
$\gamma(t)$	*The interest rate at time t*
$\delta(t)$	*The dividend at time t*
$\varepsilon(t)$	*The noise at time t*
$\theta(t)$	*The estimated mean at time t*
σ	*The standard deviation*
$\tau(t,n)$	*The relative transaction cost*

Abbreviations and Acronyms

ABM	*Agent-Based Modelling*
AIC	*Akaike's Information Criterion*
AR	*Auto Regressive*
ARCH	*Auto Regressive Conditional Heteroskedasticity*
ARMA	*Auto Regressive Moving Average*
CD	*Certificates of Deposit*
EMA	*Exponential Moving Average*
EWMA	*Exponentially Weighted Moving Average*
MA	*Moving Average*

MAE	*Mean Absolute Error*
MPT	*Modern Portfolio Theory*
MSE	*Mean Squared Error*
NASDAQ	*National Association of Securities Dealers Automated Quotations*
RSI	*Relative Strength Index*
RW	*Random Walk*
SE	*Standard Error*
STC	*Stochastic Oscillator*
TF	*Trend Following*
TR	*Trading Rule*

Contents

1

Introduction

1.1. Research Area and Problem's Actuality

The optimal financial investment (Portfolio) problem, including the forecasting and market models, was investigated by leading financial organizations and scientists. This problem is important also for small investors, who want to invest their own capital to save or enlarge it. Special attention was given to financial market analysis. A number of Nobel prizes shows the scientific recognition of this field.

The aims of most of this work are forecasting, portfolio optimization, risk minimization, and capital distribution. In some financial market research, the market prediction and portfolio optimization were regarded together. However, in most of the financial market investigations, forecast and investment problems were carried out separately. Also an important part of financial market analysis is the behavior of market's participants. There are different assumptions in this question: some scientists say that it is rational and others that it is irrational. It is a very important question, because it can explain many processes of financial market.

Effective approach of financial market investigation is the creation of its model. There are many types of models, which simulates financial (stock) market or its part: stock market games, market simulators, forecast models, and tools for market process analysis.

The financial market simulators are developed to satisfy the needs of small individual investors. The examples are the StockTrak global portfolio simulator and MarketWatch, a virtual stock exchange. Some banks offer their own investment simulators such as the Barclays "Fantasy Investment Game". 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 the computing algorithms remains unknown. Therefore, 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. Let us to mention just some examples. In (Ramanauskas and Rutkauskas, 2009) an artificial stock market by learning agents is considered.

In (Raudys and Raudys, 2011, 2012) the decisions of portfolio management were regarded in the context of artificial intelligence. In (Mockus, 2002; Raudys and Mockus, 1999; Mockus, 2012) the preliminary investigation of the virtual single stock market is discussed.

The results of the existing research helped to initiate this work modelling the stock exchange in the multi-stock financial market. Research object of this work is the development the new stock exchange model PORTFOLIO and the experimental investigation of different investment theories and strategies by this model.

1.2. Objectives and Tasks

The objective of this work is to provide a flexible, easily adaptable stock exchange model designed for the needs of individual users in the context of utility theory.

To achieve the objective, the following tasks were regarded:

1. Analysis of existing stock exchange and market models.
2. Analysis of stock price forecasting methods.
3. Analysis of portfolio optimization and trading strategies.
4. Analysis of real stock market trading strategies.
5. Development of main elements of stock exchange models, such as investors, banks, virtual stock price generators, interface to historical prices, and interconnection schemes.
6. Investigation of the price prediction algorithms.
7. Development of different short time investment strategies reflecting real practice.
8. Development of longer time trading strategies by extending short time strategies and by adding strategies based on the well-known investment theories such as the Sharp Ratio and the Markowitz Modern Portfolio Theory (MPT).
9. Performing experiments with virtual stock prices.
10. Performing experiments with historical stock prices.

The new element of the PORTFOLIO model is the investment optimization in the space of investment strategies and trading rules; both short term and longer term. The objective of the PORTFOLIO virtual part is not forecasting, but simulation of financial time series that are affected by subjective predictions of the investors. The purpose of the model is to explore the relationship between the real data and the theoretical model and to investigate what other results can be obtained using this simple model.

The new and unexpected result of experiments using the PORTFOLIO model is the observation that the minimal price prediction errors do not necessarily provide the maximal profits.

1.3. Research Methodology

Developing the new model, we use traditional prediction and investment theories and some observations of real life situations. In particular, the autoregressive models AR (p) and AR-ABS(p) are used for next day price prediction. Here parameter p (auto regression coefficient) defines a length of memory (shows how of many of previous values are used for the prediction).

In this implementation of the model, p values from 1 to 9 can be used. In addition, the Random Walk (RW) model is considered. So, 19 simple next day price forecast models can be compared. Preliminary experiments show that more complicated prediction models do not change the results significantly.

Ten different trading rules are applied for simulation of investors' behavior including four short time trading rules and six longer time ones. By combining various forecast method and trading rules we may generate 190 different investment strategies to be used by investor.

In this research, a subset of 80 investment strategies were selected by reducing the number of different $p = 1{,}3{,}6{,}9$ and performing the RW investigation separately. Virtual data was averaged by 100 tests. Historical data is of different times representing different economic conditions of approximately 360 working days each.

1.4. Scientific Novelty

1. There are many financial market models, but just a few stock exchange models. The well-known financial market models simulate interactions of independent agents trading a single stock. In contrast, the proposed

model simulates the work of stock exchange trading many different stocks.

2. New features of the proposed model:
    a) optimization in the space of investment strategies;
    b) implementation of both the real and virtual stock market in the single model;
    c) possibility of analysis of results (price prediction errors and profits) of using various trading rules and forecasting models by real and virtual data.

1.5. Results of Practical Importance

The model presents a possibility to test different investment theories and strategies using both the virtual and historical data. The model was used for graduate studies in optimization and financial markets.

1.6. Defended Claims

The PORTFOLIO model can be used to explore the relationship between the real data and theoretical assumptions and to investigate what other theoretical and practical results can be obtained using the simple stock exchange model.

The new and unexpected result of experiments using the PORTFOLIO model is the observation that minimal standard statistical stock price prediction errors do not necessarily provide the maximal profits. This result can be tested and verified independently without special skills and equipment, all the experimental conditions are defined and reproducible.

1.7. Approbation and Publications of the Research

The main results of the dissertation were published in four articles in the periodical scientific publications. The main results of the work have been presented and discussed at nine national and international conferences.

International conferences

1. The 9th International Conference Computer Data Analysis and Modelling: Complex Stochastic Data and Systems, September 7-11, 2010, Minsk, Belarus.
2. Special Workshop of Stochastic Programming Community "Stochastic Programming for Implementation and Advanced Applications" (STOPROG-2012), July 3-6, 2012, Neringa, Lithuania.
3. The 25th European Conference on Operational Research (EURO-2012), July 8-11, 2012, Vilnius, Lithuania.

Regional conferences

1. 1-oji jaunųjų mokslininkų konferencija „Fizinių ir technologijos mokslų tarpdalykiniai tyrimai", Vilnius: LMA, 2011 m. vasario 8 d.
2. 3-iasis tarptautinis seminaras "Duomenų analizės metodai programų sistemoms", Druskininkai: VU MII, 2011 m. gruodžio 1-3 d.
3. 2-oji jaunųjų mokslininkų konferencija Fizinių ir technologijos mokslų tarpdalykiniai tyrimai, Vilnius: LMA, 2012 m. vasario 14 d.
4. Lietuvos matematikų draugijos 53-oji konferencija, Klaipėda: KU, 2012 m. birželio 11-12 d.
5. 3-oji jaunųjų mokslininkų konferencija Fizinių ir technologijos mokslų tarpdalykiniai tyrimai, Vilnius: LMA, 2013 m. vasario 12 d.
6. 16-oji mokslinė kompiuterininkų konferencija „Kompiuterininkų dienos 2013", Šiauliai: ŠU, 2013 m. rugsėjo 19–21 d.

1.8. Outline of the Dissertation

The dissertation consists of 5 chapters, references and appendices. The total scope of the dissertation without appendices – 122 pages containing 231 formulas, 45 figures and 16 tables.

Chapter 1 (Introduction) presents a short description of the research context and challenges, describes the problem, the object of research, the tasks and objective of the dissertation, the methodology of research, the scientific novelty, the practical significance propositions and approbation of obtained results.

In Chapter 2, an overview of similar works is given. Detailed information about stock models, prediction methods, portfolio and trading is presented.

In Chapter 3, the diagram of the model and all the mathematical formulas describing the algorithms are presented.

Chapter 4 provides the results of experimental results.

Chapter 5 (Conclusions) presents the concluding remarks of the dissertation.

2

Financial Market Models and Simulators

Stock market is a mechanism, which set the relationship between corporations and individuals in need of funding and legal entities and individuals who can provide them with conditions. In other words, stock market gives opportunity to accumulate a capital for companies and to earn an income for investors. Stock market is a place to issue and trade shares through either exchanges or over-the-counter markets.

Also known as the equity market, it is one of the most vital areas of a market economy as it provides companies with access to capital and investors with a slice of ownership in the company and the potential of gains based on the company's future performance.

A stock exchange is a form of exchange, which provides services for stockbrokers and traders to trade stocks, bonds, and other securities. Stock exchanges also provide facilities for issue and redemption of securities and other financial instruments, and capital events including the payment of income and dividends. Securities traded on a stock exchange include shares issued by companies, unit trusts, derivatives, pooled investment products and bonds.

The aim of stock exchange models is to cover main stock exchange principals, its participants and processes between them. Developer attempts to simulate some or all features of a live stock market. These models help to understand real stock exchange principals, simulate its work. Models can be presented as computer programs or systems.

2.1. Models, Simulators and Games

After analysing stock exchange and stock market models and software, which simulates financial process, it can be divided in four main groups:

- stock market games;
- stock market models;
- stock exchange simulators;
- stock exchange models.

Stock exchange game or stock market game model simulates only stock market features such as stock prices, dividends, transaction costs, but not traders (customers). Often these models give opportunity for investors to learn by investing virtual money. Investors play in virtual stock casino with real market condition.

Examples of these models are MarketWatch (MarketWatch, 2012), StockTrak (StockTrak, 2012). Though it is possible to get some statistic data for analysis of market and its participants from these models, the main purposes of them are investment learning for new trader, understating of market dynamics, testing of price predictions and investment strategies without any risk of money losses.

In financial market models, a different approach is used. In these models, not only all market features are simulated, but also they used to simulate traders or investors behavior. To simulate behavior of trades artificial intelligent components, called agents, are used. There is so-called agent based models, where simultaneous operations and interactions of multiple agents is simulated. Agent-Based Modelling (ABM) is a method of simulation or modelling, which examines behavior of decentralized agents and how this behavior determines the behavior of the system as a whole. In contrast to the system dynamics, analyst determines the behavior of agents on the individual level, and the global behavior arises because of the activity of multiple agents (modelling „from down to up“). ABM assumes direct interaction between the agents.

The stock exchange model is used to investigate market, to create market hypotheses, to give its processes explanations, prove or disprove some market theories.

2.2. Stock Price Prediction

First step of the trader is to make his stock price or stock value prediction. There are different approaches. We discuss some of them.

2.2.1. Dividends as Main Stock Value

In artificial stock market model of T. Ramanauskas and A. V. Rutkauskas (Ramanauskas and Rutkauskas, 2009) the stock dividends is a main its value indicator. Under their theory a lot of traders make their own forecasts or fundamental market price analysis and they affect market. But before trading it’s useful to see company’s financial books, because there they could see real company’s financial value. Also these authors consider, that some traders make their decisions on their own believes in stock value. This also reflected in stock’s current price. In spite of this, authors mean, that main index of stock value is its dividends. Because of that, in their model all trading agents make their decisions based on dividends dynamic. At first, agents determine the basic reference point for their dividends forecast. Here for calculations method of Exponentially Weighted Moving Average (EWMA) was used. Later in dividends forecasts calculation adjustment factor or coefficient used. These adjustment factors changes after agents explore and exploit their accumulated experience, with the long-term aim to minimize squared forecast errors.

On the next step of stock value calculation, agents estimates their stock reservation price, which includes dividends calculation and adjustment coefficient. This price is used by an agent to make his decision: to buy or to sell stock.

Agents start with determining basic reference points for their dividend forecasts. EWMA of realised dividend payouts can be calculated as follows:

$$d_{i,y}^{\text{EWMA}} = \lambda_1 \cdot d_y + (1 - \lambda_1) d_{i,y-1}^{\text{EWMA}}. \tag{2.1}$$

Here d_y denotes dividends paid out in period y (year) and λ_1 is the arbitrary smoothing factor. This factor is the same for all agents and its value always between 0 and 1.

The *n*-period dividend forecast is given by the following equation:

$$E(d_{i,y+n}) = d_{i,y}^{\text{EWMA}} \cdot a_{i,y}^{\text{div}}, \tag{2.2}$$

where $a_{i,y}^{\text{div}}$ is agent i's dividend adjustment factor. These adjustment factors are gradually changed as agents explore and exploit their accumulated experience, with the long-term aim to minimize squared forecast errors.

Authors assume that agents' behavior is driven by reinforcement learning since these learning algorithms borrowed from the machine learning literature seem to be conceptually suitable for modelling investor behavior.

Individual forecasts for periods $y + 1, \dots, y + n$ formed in periods $y - n + 1, \dots, y$, respectively, are stored in the program and used for determining individual estimates of the fundamental stock value.

2.2.2. Stock Market Forecasting Using Machine Learning Algorithms

In (Shen, Jiang and Zhang, 2012) authors proposed the use of global stock data in associate with data of other financial products as the input features to machine learning algorithms such as support vector machine (SVM) and reinforcement learning.

In this project, authors tries to predict the trend of stock market (either increase or decrease). They assume that the change of a feature over time is more important than the absolute value of each feature. Here feature *i* at time *t* defined as $x_i(t)$, where $i \in \{1,2,\dots,16\}$. The feature matrix is given by

$$F = (X_1, X_2, \dots, X_n)^T, \tag{2.3}$$

where

$$X_t = (x_1(t), x_2(t), \dots, x_{16}(t)). \tag{2.4}$$

The new feature which is the difference between two daily prices can be calculated by

$$\nabla \delta x_i(t) = x_i(t) - x_i(t-\delta), \tag{2.5}$$

$$\nabla_\delta X(t) = x(t) - X(t-\delta) = (\nabla_\delta x_1(t), \nabla_\delta x_2(t), \dots, \nabla_\delta x_{16}(t))^T, \tag{2.6}$$

$$\nabla_\delta F = (\nabla_\delta X(\delta+1), \nabla_\delta X(\delta+2), \dots, \nabla_\delta X(n)). \tag{2.7}$$

Here due markets basic and their value difference calculated differentials can vary in a wide range. To make them comparable, the features are normalized as following:

$$N(\nabla_\delta x_i(t)) = \frac{x_i(t) - x_i(t-\delta)}{x_i(t-\delta)},$$
$$N(\Delta_\delta X(t)) = \Big(N(\nabla_\delta x_1(t)), \dots, N(\nabla_\delta x_{16}(t))\Big)^T, \tag{2.8}$$
$$N(\nabla_\delta(F)) = \Big(N(\nabla_\delta X(\delta+1)), \dots, N(\nabla_\delta X(n))\Big)^T,$$

and the normalization can be implemented as:

$$\text{normal}(X(t)) = \frac{N(\nabla_\delta X(t))}{|N(\nabla_\delta X(t))|}. \tag{2.9}$$

It is assumed that performance of stock market predictor mostly depends on correlation between the date used for training and the current data for prediction. In other words, if the trend of stock price is always an extension of previous, the accuracy of prediction should be fairly high. To select input features with high temporal correlation, authors calculated the autocorrelation and cross-correlation of different market trends (increase or decrease).

2.2.3. Time-Series Forecasting Algorithms

In (Zuo and Kita, 2011) authors analize time-series forecast algorithms, using them in stock price forecasting. In this chapter, we consider the definitions of time-series prediction algorithms given by these authors.

2.2.3.1 Auto Regressive (AR) Model

The notation r_t denotes the price earnings ratio (P/E ratio) of the stock at time *t*. In AR model AR(*p*), the P/E ratio r_t is approximated with the previous P/E ratio r_{t-i} $(i = 1, \dots, p)$ and the error term u_t as follows:

$$r_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i} + u_t\,, \tag{2.10}$$

where α_i $(i = 0, \dots, p)$ is the model parameter. The error term u_t is a random variable from the normal distribution centered at 0 with standard deviation equal to σ^2.

2.2.3.2 Moving Average (MA) Model

In the MA model MA(*q*), the P/E ratio r_t is approximated with the previous error term u_{t-j} $(j = 1, \dots, q)$ as follows:

$$r_t = \beta_0 + \sum_{j=1}^{q} \beta_j u_{t-j} + u_t\,, \tag{2.11}$$

where β_j $(j = 0, \dots, q)$ is the model parameter.

2.2.3.3 Auto Regressive Moving Average (ARMA) Model

The ARMA model is the combinational model of AR and MA models. In the ARMA model ARMA(p, q), the P/E ratio r_t is approximated as follows:

$$r_t = \sum_{i=1}^{p} \alpha_i r_{t-i} + \sum_{j=1}^{q} \beta_j u_{t-j} + u_t \,. \tag{2.12}$$

2.2.3.4 Auto Regressive Conditional Heteroskedasticity (ARCH) Model

In the ARCH model ARCH (p, q), the P/E ratio r_t at time t is approximated as follows:

$$r_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i} + u_t \,. \tag{2.13}$$

The error term u_t is as:

$$u_t = \sigma_t z_t, \tag{2.14}$$

where $\sigma_t > 0$ and the function z_t is a random variable from the normal distribution centered at 0 with standard deviation equal to 1.

The volatility σ_t^2 is approximated by the following expression:

$$\sigma_t^2 = \beta_0 + \sum_{j=1}^{q} \beta_j u_{t-j}^2 \,. \tag{2.15}$$

2.2.3.5 Determination of Model Parameters

In each model, the model parameters p and q were fixed integers. Akaike's Information Criterion (AIC) is estimated in all cases. The parameters p and q for maximum AIC are adopted.

The AIC is given as follows:

$$\text{AIC} = \ln\hat{\sigma}^2 + \frac{2(p+q)}{T}, \tag{2.16}$$

where $\hat{\sigma}$ is the volatility estimated from the model error $\epsilon_1, \epsilon_2, \dots, \epsilon_T$. T in this equation denotes time period.

Authors used these algorithms for NIKKEI stock index and TM (Toyota

Motors) stock prediction. No discussion of global optimization issues was presented which is needed for optimization of models' parameters.

2.3. Trading Strategies and Portfolio Problem

2.3.1. Modern Portfolio Theory (MPT)

MPT (Marling and Emanuelsson, 2012) was developed for portfolio selection and portfolio optimization. It provides the foundation for MPT as a mathematical problem.

The return R_t of a portfolio at time *t* is defined by the following formula:

$$R_t = \frac{T_t}{T_{t-1}} - 1. \tag{2.17}$$

where T_t is the total value of the portfolio at time t.

Markowitz portfolio theory provides a method to analyse portfolio quality based on the means and the variances of the returns of the assets contained in the portfolio. An investor is supposed to be risk-averse hence he/she wants a small variance of the return (i.e. a small risk) and a high expected return.

Consider a portfolio with *n* different assets where asset number *i* will give the return R_i. Let μ_i and σ_i^2 be the corresponding mean and variance and let $\sigma_{i,j}$ be the covariance between R_i and R_j. Suppose the relative amount of the value of the portfolio invested in asset *i* is x_i. If *R* is the return of the whole portfolio, then:

$$\mu = E[R] = \sum_{i=1}^{n} \mu_i x_i, \tag{2.18}$$

$$\sigma^2 = \mathrm{var}[R] = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{i,j} x_i x_j\,, \tag{2.19}$$

$$\sum_{i=1}^{n} x_i = 1\,, \tag{2.20}$$

$$x_i \geq 0, i = 1,2, \dots, n. \tag{2.21}$$

For different choices of $x_1, \dots, x_n$ the investor will get different μ and σ^2. A set of all possible pairs (σ^2, μ) is called the attainable set. Those (σ^2, μ) with minimum σ^2 for a given μ and maximum μ for a given σ^2 are called the efficient set (or efficient frontier). Since an investor wants a high profit and a small risk he/she wants to maximize μ and minimize σ^2 and therefore he/she should choose the portfolio (σ^2, μ) which is in the efficient set. In Figure 1, the attainable set is the interior of the ellipse and the efficient set is the upper left part of the boundary.

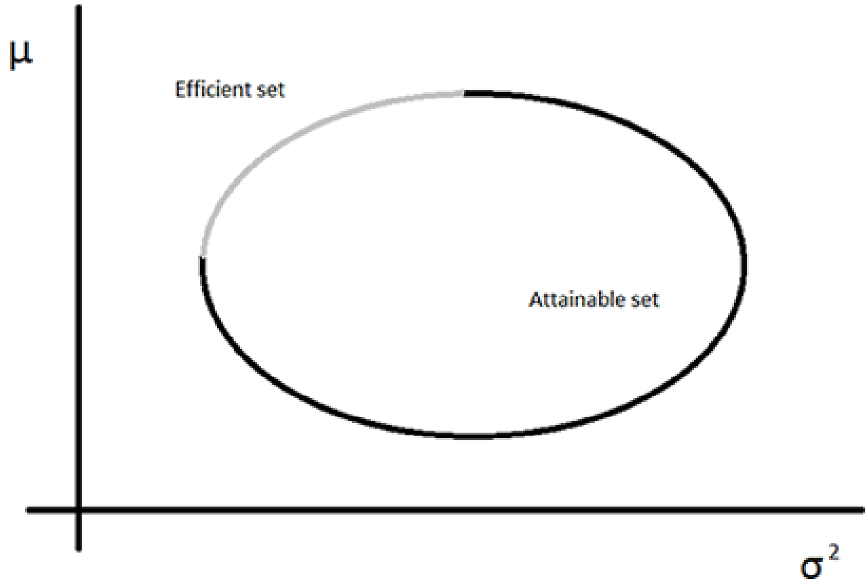

Fig. 1 *The efficient set in the* $\boldsymbol{\mu\sigma^2}$ *plane*

2.3.2. Trend Following (TF) Algorithms

Another stock trading method is Trend Following (TF) (Fong, Si and Tai, 2012). It is a trading method in response to the real-time market situation. The trading decisions are made according to observed market trend. If the trend is identified, it activates the trading rules and adheres rigidly to the rules until the next prominent trend is identified. TF does not guarantee profit every time, but nonetheless in a long-term period it may probably profit by obtaining more gains than loses.

The nature of TF makes it as an ideal ingredient in implementing a decision-making component in automated trading software where human intervention is not required. This method was used in the software of trading algorithms for many years. In the next section we present three TF algorithms. Examples are in pseudo-code.

2.3.2.1 Static TF Algorithm

The algorithm finds the trend, identifies the trade signals and trade on that signals until the end of this trend. It is assumed that trend is more likely to continue than to reverse.

In the Static TF algorithm, two constants are used as the two comparison marks when substantial change in the trend would trigger the trading system to open or close a position accordingly. These constants are defined as P and Q, where P is the amount of up-trend required for opening a position, and Q is the amount of opposite trend required to close this position.

In reality market price does not move in a straight line. It is therefore impractical to apply the P and Q rules directly on the trend T, because the frequent fluctuation will generate too many signals of trading actions. An Exponential Moving Average (EMA) algorithm is used to smooth out this fluctuation, which is as follow:

$$\mathrm{EMA}_{(t)} = \left(\mathrm{price}_{(t)} - \mathrm{EMA}_{(t-1)} \times \frac{2}{n+1}\right) + \mathrm{EMA}_{(t-1)}, \tag{2.22}$$

Algorithm 1. Pseudo-codes of the Static TF algorithm.

```
Repeat until end of market
Compute EMA(T)
If no position opened
If EMA(T) >= P
If trend is going up
Open a long position
```

Else if trend is going down

Open a short position

Else if any position is opened

If EMA(¬T) >= Q

Close position

If end of market

Close all opened position

Here EMA(T) is Exponential Moving Average of the real time market price trend, and EMA(¬T) is the reversion (opposite direction) of the trend counting from the highest (or lowest) point of this trend.

2.3.2.2 Dynamic TF Algorithm

In the dynamic TF algorithm, *P* and *Q* are variables, instead of static constants, and their values change adaptively to the current market trends. Based upon this initial concept of trading algorithm, dynamic TF algorithm is introduced with incorporation of technical analysis concept. Technical analysis makes trade decision through technical indicators such as Relative Strength Index (RSI), Stochastic Oscillator (STC) and EMA. These indicators are changing dynamically according to the market situation. By adopting one or more of these indicators and by studying how they react to the market, some rules can be formed that are able to inherit this dynamic nature. By following these rules during trade session, we update the trade parameters *P* and *Q* with the latest dynamic values attribute.

There are hundreds of indicators in use today, but not all are tested to be reliable. Experiments have been conducted to try out many popular ones and RSI is found to be the best for TF. RSI compares the magnitude of underlying recent gains of an asset to the magnitude of its recent losses, and normalized to a number that ranges from zero to 100.

$$\mathrm{RSI}_{(t)} = 100 - \frac{100}{1 + \mathrm{RS}}, \tag{2.23}$$

$$\mathrm{RS} = \frac{\mathrm{AU}(t)}{\mathrm{AD}(t)}, \tag{2.24}$$

$$\mathrm{AU}(t) = \frac{\mathrm{Up}(t) + \mathrm{Up}(t-1) + \cdots + \mathrm{Up}(t-n+1)}{n}, \tag{2.25}$$

$$\mathrm{AD}(t) = \frac{\mathrm{Down}(t) + \mathrm{Down}(t-1) + \cdots + \mathrm{Down}(t-n+1)}{n}, \tag{2.26}$$

where AU is average price upward movement in *n* periods, AD is average price decline in *n* periods, *t* is the time, *n* is the number of RSI periods usually 14. STC is a momentum indicator that shows the location of the current close over a number of periods

$$\%\mathrm{K}(t) = 100 \times \frac{\mathrm{Close}(t) - \mathrm{LL}(n)}{\mathrm{HH}(n) - \mathrm{LL}(n)}, \tag{2.27}$$

$$\%\mathrm{D}(t) = \mathrm{EMA}(\%\mathrm{K}(t))(m), \tag{2.28}$$

where HH is highest high in *n* periods, LL is lowest low in *n* periods, *n* is number of STC periods, *m* is number of periods of EMA that applied on %K.

Algorithm 2. Pseudo-codes of the dynamic TF algorithm.

Repeat until end of market

Compute RSI(*t*) and RSI(EMA(*t*))

If price is advancing:

If RSI(*t*) > EMA(*t*) and 40 < EMA(*t*) > 60

If no position has been opened

Open a long position

Else if short position has been opened

Close out short position

Else if price is declining:

If RSI(*t*) < EMA(*t*) and 40 < EMA(t) > 60

If no position has been opened

Open a short position

Else if long position has been opened
Close out long position
If end of market
Close all opened positions

2.3.2.3 Fuzzy TF algorithm

The static and dynamic TF algorithms described in previous section are designed to make trading decisions based on criteria, which are formulated in classical binary logic. In this section, we consider TF algorithms based on Fuzzy logic (Zadeh, 1973). jFuzzyLogic (jFuzzyLogic, 2012) was used to develop a fuzzy inference system. Based on our experience with previous TF algorithms, we define three membership functions for input and output variables.

The fuzzy inference engine accepts RSI and momentum indicator (MTM) as input and produces recommendations on whether or not to take a position (POS) as output.

MTM is an oscillator type indicator used to detect overbought and oversold conditions and to perform as a gauge indicating the strength of the current trend. MTM calculations are either positive or negative and fluctuate around a zero line:

$$\mathrm{MTM}(t) = C_{(t)} - C_{(t-n)}, \tag{2.29}$$

where $C_{(t)}$ is the closing price, *n* is the number of MTM periods.

1. IF RSI IS whipsaw OR MTM IS whipsaw THEN POS IS doNothing.
2. IF RSI IS overSold AND MTM IS long THEN POS IS goLong.
3. IF RSI IS overBought AND MTM IS short THEN POS IS goShort.
4. IF RSI IS overSold AND MTM IS short THEN POS IS goShort.
5. IF RSI IS overBought AND MTM IS long THEN POS IS goLong.

Whipsaw is a condition where a security’s price heads in one direction is followed quickly by a movement in the opposite direction. Whipsaw pattern for RSI can be considered as a neural signal in terms of the velocity and magnitude of directional price movements. The security is considered to be in overbought territory when RSI is above 70 and considered to be over sold when RSI is below 30. Momentum shows the difference between today's closing price and the closing price of *n* days ago:

$$\text{momentum} = \text{close}_{\text{today}} - \text{close}_{n \text{ days ago}}. \tag{2.30}$$

2.4. Existing Stock Market Models and Virtual Stock Markets

2.4.1. MarketWatch – Free Stock Market Game

The MarketWatch virtual stock game is a competition game under real market rules, where customer invests his fixed virtual money budget into stocks. In this game a task is to maximize profit and to win between many players. Games environment uses real stock prices, but other things are virtual. Player doesn't risk with his own funds, but can learn to invest like he would buy real stock. No stock market model is described.

This stock market game is intended as a tool to learn how to analyze data. No formal stock market and stock exchange models are applied.

2.4.2. NASDAQ Market Simulation

Another stock market model is NASDAQ market simulation developed by Vince Darley and Alexander Outkin (Darley and Outkin, 2004). This model is stock market model, where stock trading occur between two customers, but not between customer and exchange.

This model based on the Glostem-Milgrom model. This model simplifies complex real market interaction. The main assumption of this model is that there are informed traders on market, who exactly knows real stock price.

Informed traders have access to additional information about the realization of a security's true value, V. In the elementary version of the Glosten-Milgrom model, the distribution of V is binomial (θ, $\underline{V}$, $\overline{V}$): with probability θ that the variable V is equal to $\overline{V}$, and with probability $1-\theta$ that it is equal to a lesser value, $\underline{V}$. The presence of traders with superior information leads to a positive bid-ask spread even when the trader is risk-neutral and makes zero expected profits. Important limitation of this model is that market has only one asset and all trades are risk neutral and are trading directly with each other. There are no transaction costs, taxes, portfolio expenses or banks in this model. Model has an auction mechanism and traders are agents, who use learning strategies. Agents learn to interact with unknown environment for long period, while trying to maximize profit.

2.4.3. Artificial Stock Market

In the Artificial stock market by Tomas Ramanauskas and Aleksandras Vytautas Rutkauskas (Ramanauskas and Rutkauskas, 2009) the dividends are used as main stock price evaluation unit. This model is on such stock market, where independent traders directly interact with each other. The model is based on interaction of heterogeneous agents whose forward-looking behavior is driven by the reinforcement learning algorithm combined with some evolutionary selection mechanism. The model is similar to the NASDAQ model, where for agent-environment interaction Q-learning algorithm is used.

2.5. Conclusions of Chapter 2

The main-stream of publications is on forecasting and portfolio selection. Unexpectedly, there are just a few publications on developing and investigation of the stock exchange models.

3

PORTFOLIO Model

In this chapter, the proposed stock exchange model PORTFOLIO is described. The algorithmic diagram and the process logic are presented. Here are all the mathematical formulas, which describe model’s basic processes and strategies.

Those strategies include prediction models and trading rules. The profits of both investors and banks are calculated. Most of the formulas are new, they describe the new elements of the model. However, some formulas describing the previous models are included too, for the consistency. The corresponding experimental results are in the fourth chapter.

The PORTFOLIO model simulates behavior of group of investors, who trade stocks in real and virtual environments. The optimization is performed on a set of investment strategies. This is the main specific feature of the PORTFOLIO model. Investors can choose one of 190 investment strategies, including ten trading rules and nineteen forecasting models. Three of these trading rules model known theoretical results, the others are new and simulates heuristics of different investors with different approaches to risk.

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

The aim of the PORTFOLIO model is not forecasting, but analyzing of stock exchange processes, verifying various market hypotheses, testing market manipulation tools and understanding the differences between the real and virtual environments.

To make flexible and easily adaptive stock exchange model, java applet technology was selected. Model's software is written with java programming language, using objective oriented methodology. The structure of software presents possibility to extend model: to add new prediction methods and trading rules. Implementation of the model as java applet allows its application by any web-browser with Java support. For 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. The description of software is in the Appendix.

3.1. Basic PORTFOLIO Scheme

In this chapter, the basic algorithmic scheme is presented first, see Figure 2. It reflects model's workflow and shows main blocks of it.

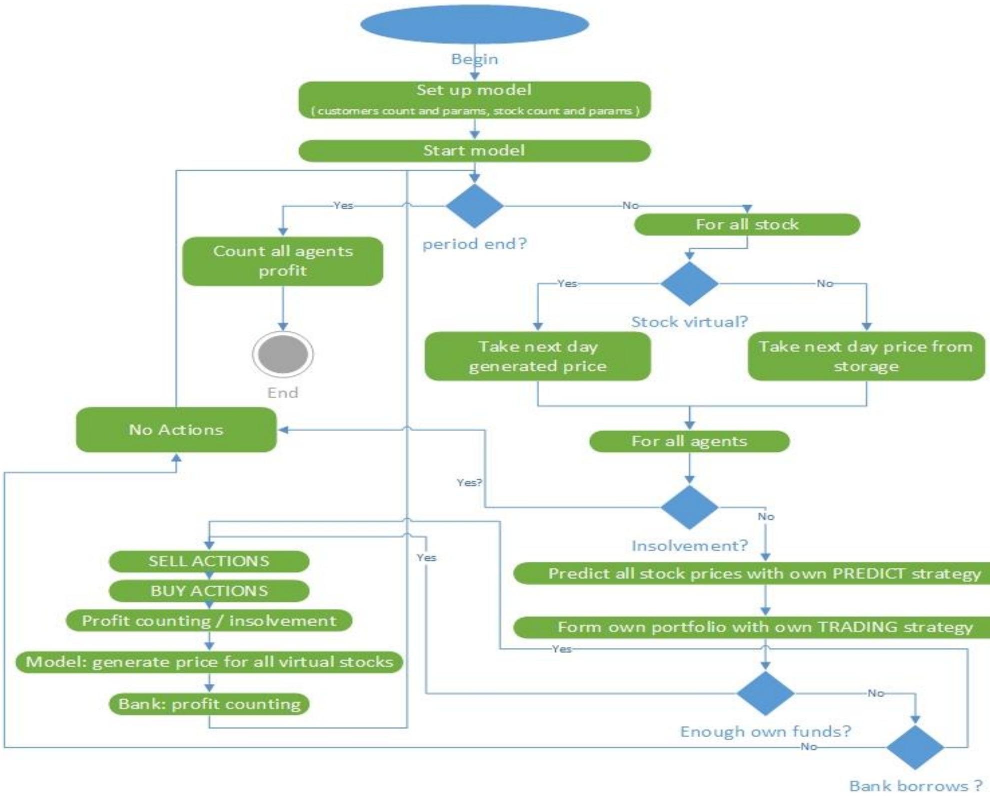

Fig. 2 *Basic PORTFOLIO scheme*

3.2. Main Models Concept

3.2.1. Basic Buying and Selling Strategies

The PORTFOLIO model simulates banks and major investors buying-selling stocks of different joint-stock companies assuming I major players ($i = 1, \ldots, I$).

A new feature of the model is simulation of multi-stock environment. This is important representing the processes of real and virtual stock exchange. So, different trading rules and different prediction modes can be investigated using the PORTFOLIO model using both the historical and virtual data generated simulating behavior of different investors.

However, we start by presenting mathematical formulation of single stock trading, for simplicity. The single-stock assumption was used in the prototype model (Mockus, 2012) and in the models describing direct interaction investors by (Darley and Outkin, 2004) and (Ramanauskas and Rutkauskas, 2009). We shall use the notations similar to those in (Mockus, 2012).

The main variables of the simplified model are as follows:

$z(t) = z(t, i)$ 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 the dividend at time t,

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

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

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

$$\beta(t, i) = \frac{z(t + 1, i) - Z(t)}{Z(t)}. \tag{3.1}$$

In the PORTFOLIO model, the investors decisions depend on the expected profitability[2] (relative profit). It is defined as the relative profit $p(t, i)$ of an investment at 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 $\gamma(t)$:

$$p(t, i) = \begin{cases} \beta(t) + \delta(t) - \gamma(t), & \text{investing borrowed money,} \\ \beta(t) + \delta(t) - \alpha(t), & \text{investing own money.} \end{cases} \tag{3.2}$$

The aim is profit, thus a customer i will buy some number $n_b(t, i) \geq n(t)$ of stocks, if profitability is greater comparing with the relative transaction cost $\tau(t, n)$; $p(t, i) > \tau(t, n)$, will sell stocks, if the relative loss (negative profitability $-p(t, i)$) is greater as compared with the transaction cost $p(t, i) < -\tau(t, n)$, and will do nothing, if $-\tau(t, n) \leq p(t, i) \leq \tau(t, n)$. Here the relative

[1] The term “actual” means simulated by PORTFOLIO.
[2] The term “profit” can define losses if negative terms prevail.

transaction cost is defined as the relation:

$$\tau(t,n) = \frac{\tau_0}{n(t)Z(t)}, \tag{3.3}$$

where τ_0 is the actual transaction cost and $n = n(t)$ is the number of transaction stocks. From equality $\tau(t,n) = p(t,i)$ it follows that a minimal number of stocks to cover transaction expenses is

$$n(t) = \frac{\tau_0}{p(t,i)Z(t)}. \tag{3.4}$$

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

$$S(t,i) =$$

$$= \begin{cases} \text{buy } n_b(t,i) \geq n(t) \text{ stocks,} & \text{if } p(t,i) \geq \tau(t,n) \text{ and } n \leq n_b^{\max}, \\ \text{sell } n_s(t,i) \geq n(t) \text{ stocks,} & \text{if } p(t,i) \leq -\tau(t,n) \text{ and } n \leq n_s^{\max}, \\ \text{wait,} & \text{if } |p(t,i)| \leq \tau(t,n^{\max}). \end{cases} \tag{3.5}$$

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

If

$$n_b(t,i) = n_b^{\max} \text{ and } n_s(t,i) = n_s^{\max}, \tag{3.6}$$

then this buying/selling strategy reflects the behavior of risk-neutral stockholders which invest all available resources if the expected profitability is higher than the transaction cost. If the expected losses are greater, then all the stocks are sold. This means that stockholders may tolerate considerable probability of losses if the expected profits are positive. This way, the maximal expected profit is provided. However, the probability to get losses instead of profits could be near to 0.5.

From expressions (3.1) and (3.2), the buying-selling strategy $S(t,i)$ in terms of stock price levels:

$$S(t,i) =$$

$$= \begin{cases} \text{buy } n_b(t,i) \geq n(t) \text{ stocks,} & \text{if } Z(t) \leq z_b(t,n,i) \text{ and } n \leq n_b^{\max}, \\ \text{sell } n_s(t,i) \geq n(t) \text{ stocks,} & \text{if } Z(t) \geq z_s(t,n,i) \text{ and } n \leq n_s^{\max}, \\ \text{wait,} & \text{otherwise.} \end{cases} \tag{3.7}$$

Here the price level of the player i to buy at least $n = n(t)$ stocks at time t is

$$z_b(t,n,i) = \frac{z(t+1,i)}{1-\delta(t)+\alpha(t)+h(t)+\tau(t,n)}. \tag{3.8}$$

The price level of the player i to sell at least $n = n(t)$ stocks at time t is

$$z_s(t,n,i) = \frac{z(t+1,i)}{1-\delta(t)+\alpha(t)+h(t)-\tau(t,n)}, \tag{3.9}$$

where $z(t+1,i)$ is the stock price predicted by the investor i at time $t+1$.

The market buying price at time t is the largest buying price of players $i = 1,\ldots,I$: $z_b(t,n) = z_b(t,n,i^{\max})$, 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 $i = 1,\ldots,I$: $z_s(t,n) = z_b(t,n,i^{\min})$, where $i^{\min} = arg\min_i z_s\,(t,n,i)$.

3.2.2. Gaussian Model for Next Day Price Generation

In virtual market next day price or actual price is generating by model. 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.

$$Z(t+1) = \begin{cases} z_b(t,n) + Z(t) + \epsilon(t+1), & \text{if } Z(t) < z_b(t,n), \\ z_s(t,n) + Z(t) + \epsilon(t+1), & \text{if } Z(t) > z_s(t,n), \\ Z(t) + \epsilon(t+1), & \text{if no deal.} \end{cases} \tag{3.10}$$

The noise is defined as truncated Gaussian random number $gaussian(0,v(i))$ with standard deviation $v(i)$ which reflects stocks volatility.

Here the noise $\varepsilon(t+1)$ is generated by the truncated Gaussian distribution with minimal values restricted by this condition

$$Z(t+1,i) \geq \rho + 3\tau_0, \tag{3.11}$$

where $\rho > 0$ is the minimal stock price "insolvency level". This inequality is implemented by ignoring those ε values that are crossing the lower limit

(3.11). That means that we accept only those noise $\varepsilon(t+1)$ values which satisfy this condition:

$$\varepsilon(t+1) = gaussian\big(0, v(i)\big) - \theta(t), \text{if } Z(t+1, i) \geq \rho + 3\tau_0. \quad (3.12)$$

The other random numbers generated by $gaussian(0, v(i))$ are simply omitted.

Inequality (3.11) is necessary to represent real stock prices but it violates conditions of the Wiener process. This means that we simulate not the genuine Wiener process but some approximation. The important difference is that expectation of truncated Gaussian noise is positive. In the expression (3.12), to eliminate this difference, we subtract the estimated mean $\theta(t)$:

$$\theta(t) = \frac{1}{t}\sum_{s=1}^{t} \varepsilon\,(s). \quad (3.13)$$

The problem is how to define $Z(t+1)$, if several buying conditions by different stockholders i are satisfied at the same time. A possible solution is to prefer the user which buying price level $z_b(t, i, l)$ is crossed by the actual price line $Z(t)$ first assuming that this user has sufficient funds.

The Gaussian distribution follows from the assumption that the noise is a sum of many independent random numbers representing the buying/selling actions of remaining small stockholders. In the other well-known approach (Wilmott, 2007), the log-normal distribution of $\varepsilon(t+1)$ is considered. The log-normal distribution follows from the assumption that the noise is a product of many random variables. In (Landauskas and Valakevicius, 2011) the technique involving Markov Chain Monte Carlo (MCMC) sampling from piece-wise-uniform distribution is proposed.

3.2.3. Market Inertia

In PORTFOLIO market, an inertia coefficient is introduced to represent the inertia of real and virtual markets. The corresponding modification of stock price calculation is as follows:

$$Z(t+1) = $$

$$= \begin{cases} (1-a)z_b(t,n) + aZ(t) + \epsilon(t+1), & \text{if } Z(t) < z_b(t,n), \\ (1-a)z_s(t,n) + aZ(t) + \epsilon(t+1), & \text{if } Z(t) > z_s(t,n), \\ Z(t) + \epsilon(t+1), & \text{if no deal.} \end{cases} \quad (3.14)$$

The modified expression includes the present stock price, too. This way we are taking into account some inertia of the stock market with large number of small shareholders. The market inertia level is defined by a multiplier $0 \le a \le 1$, where $a = 0$ means no inertia and $a = 1$ describes maximal inertia (no market reaction to the last deal). In addition, it regards the situations when the buying price of the offer is higher than the market price and the selling price is lower than the market price at this moment.

In PORTFOLIO this coefficient can be defined by user. Also it is supposed that inertia is more important just after some new shares are introduced in the market. So, in the present software the parameter $a = 1.0$ if $t \le 20$ by default. After this time, we control the market inertia by setting the parameter a.

3.2.4. Buying-Selling Price

The market buying price at time t is the largest buying price of players $i = 1, \ldots, I$: $z_b(t,n) = z_b(t,n,i^{\max})$, 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 $i = 1, \ldots, I$: $z_s(t,n) = z_s(t,n,i^{\min})$, 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

$$N(t+1,i) = \begin{cases} N(t,i) + n_b(t,n,i), & \text{if } Z(t) < z_b(t,n), \\ N(t,i) - n_s(t,n,i), & \text{if } Z(t) > z_s(t,n), \\ N(t,i), & \text{if no deal.} \end{cases} \quad (3.15)$$

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

3.2.5. 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 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 this recurrent expression:

$$C_0(t,i) = C_0(t-1,i) - (N(t,i) - N(t-1,i))\,Z(t), \tag{3.16}$$

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 investors own money $C_0(t,i)$ and the funds $b(t,i)$ borrowed at moment t. The borrowed sum of the stockholder i accumulated at time t is

$$B(t,i) = \sum_{s=1}^{t} b\,(s,i). \tag{3.17}$$

The symbol $b(t,i)$ shows what the user i borrows at moment $s = t$:

$$b(t,i) = \\ = \begin{cases} -C_0(t,i), & \text{if } -L(t,i) \le C_0(t,i) < 0, \\ 0, & \text{if } 0 \le C_0(t,i), \\ \text{insolvent at moment } t = t_i^*, & \text{if } -L(t,i) > C_0(t,i) - B_{\text{sum}}(t,i) + N(t,i)\,Z(t). \end{cases} \tag{3.18}$$

Expression (3.14) is for long-term loans where frequent transactions are not economical or restricted by contracts. The advantage is lower interest rate $\gamma(t)$.

$$b(t,i) = \\ = \begin{cases} -C_0(t,i), & \text{if } -L(t,i) \le C_0(t,i) < 0, \\ -C_0(t,i), & \text{if } 0 \le C_0(t,i) < B(t,i), \\ 0, & \text{if } B(t,i) \le C_0(t,i), \\ \text{insolvent at moment } t = t_i^*, & \text{if } -L(t,i) > C_0(t,i) - B_{\text{sum}}(t,i) + N(t,i)\,Z(t). \end{cases} \tag{3.19}$$

According to the second line in expression (3.18), the user i "borrows" a negative sum $b(t,i) = -C_0(t,i)$ if $0 < C_0(t,i) \le B(t,i)$, which means that the user pays back a part $b(t,i)$ of the loan $B(t,i)$ using available funds

$C_0(t, i)$. This expression is for short-term loans with possibility of frequent transactions. The disadvantage is a higher interest rate $\gamma(t)$.

The general borrowing expenses are

$$B_{\text{sum}}(t, i) = B(t, i) + \sum_{s=1}^{t} B\,(s, i)\,\gamma(s, i), \tag{3.20}$$

where the first term denotes the loan accumulated at time T and the second term shows the interest.

An investor (stockholder) gets a profit as the difference between the income from selling and buying stocks $D(t, i)$ and expenses for borrowing funds $B_{\text{sum}}(t, i)$:

$$U(t, i) = D(t, i) - B_{\text{sum}}(t, i), \tag{3.21}$$

where

$$D(t, i) = N(t, i)Z(t) - N(0, i)Z(0). \tag{3.22}$$

The funds available for the investor i at time t are

$$C(t, i) = C_0(t, i) + L(t, i) - B_{\text{sum}}(t, i). \tag{3.23}$$

An investor is trying to maximize gains by borrowing money to invest in shares that appreciate more than what it costs him by way of interest. It means leveraging shares for an investment.

The number of stocks $n_b(t)$ to buy at the time t is restricted by the following inequality:

$$n(t) \leq n_b(t, i) \leq \frac{C(t, i)}{Z(t)}. \tag{3.24}$$

Here the first part of the inequality restricts transaction costs. According to expression (3.18), the stockholder will be insolvent at the time $t = t_i^*$ if the loan exceeds the assets

$$B_{\text{sum}}(t_i, i) > C_0(t_i, i) + L(t_i, i) + N(t, i)\,Z(t), \tag{3.25}$$

since there will not be enough money to pay back all the borrowing expenses $B_{\text{sum}}(t_i^*, i)$. This can happen without buying additional stocks, because the interest $B_{\text{sum}}(t, i)$ accumulates automatically.

Considering multi-level operations, we shall define additional restrictions (3.40) on the number of stocks $n_b(t)$.

3.2.6. Bank Profit

It follows from (3.25) that the bank losses at time t_i^* are

$$B_{\text{loss}}(t_i^*, i) = B_{\text{sum}}(t_i^*, i) - C_0(t_i^*, i) - N(t_t^*, i)\, Z(t_i^*). \tag{3.26}$$

The total bank losses accumulated at time $t \geq \max_i t_i^*$ are

$$B_{\text{loss}}(t) = \sum_i B_{\text{loss}}\,(t_i^*, i). \tag{3.27}$$

The bank income:

$$D(t) = \sum_{s=1}^{t} \sum_{i=1}^{I} B\,(s, i)\, \gamma(s, i). \tag{3.28}$$

The bank profit:

$$U(t) = D(t) - B_{\text{loss}}(t). \tag{3.29}$$

3.2.7. Multi-Level Operations

In the opinion of some professional brokers we have interviewed, to represent risk-aware stockholders one needs at least three buying profitability levels $p_b(t, i, l)$, $l = 1{,}2{,}3$, where

$$p_b(t, i, l+1) > p_b(t, i, l), p_b(t, i, 1) = \tau(t), \tag{3.30}$$

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

$$p_s(t, i, l+1) < p_s(t, i, l), p_s(t, i, 1) = -\tau(t), p_b(t, i, l) > p_s(t, i, 1). \tag{3.31}$$

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.

Thus, the number of stocks to buy at time t and the profitability level $l = 3$ is as follows:

$$n_b(t,i,3) = \text{int}\left(\frac{C(t,i)}{Z(t)}\right), \text{if } p(t,i) \geq p_b(t,i,3). \tag{3.32}$$

The number of stocks to buy at time t at the profitability level $l = 2$:

$$n_b(t,i,2) = \text{int}\left(\frac{C(t,i)}{2Z(t)}\right), \text{if } p_b(t,i,2) \leq p(t,i) < p_b(t,i,3). \tag{3.33}$$

The number of stocks to buy at time t at the profitability level $l = 1$:

$$n_b(t,i,1) = 1, \text{if } p(t,i,1) \leq p(t,i) < p_b(t,i,2). \tag{3.34}$$

We do not sell, if the maximal expected losses are less than the transaction cost $N(t,i)Z(t)p(t,i) < -\tau(t)$, where $N(t,i)$ is the number of stocks available at time t.

This buying-selling strategy approximately describes the risk-averse stockholders since they invest larger sums if the probability of losses is smaller.

The feasible number of stocks to be sold at time t and the selling profitability level $l = 3$ is

$$n_s(t,i,3) = N(t,i), \text{if } p(t,i) \leq p_s(t,i,3). \tag{3.35}$$

The number of stocks to be sold at time t and the selling profitability level $l = 2$ are

$$n_s(t,i,2) = \frac{N(t,i)}{2}, \text{if } p_s(p(t,i,3) > p(t,i) \leq p_s(t,i,2), \tag{3.36}$$

and the number of stocks to be sold at time t and the selling profitability level $l = 1$ are

$$n_s(t,i,1) = 1, \text{if } p_s(t,i,2) > p(t,i) \leq \ p_s(t,i,1). \tag{3.37}$$

Here

$$n_b(t,i,l) \leq n_b(t,i,l+1), l = 1,2,3, n_b(t,i,3) = \text{int}\left(\frac{C(t,i)}{Z(t)}\right),$$
$$n_s(t,i,l) \leq n_s(t,i,l+1), l = 1,2,3, n_s(t,i,3) = N(t,i). \tag{3.38}$$

The general buying-selling strategy $S_0(l,i)$ of the investor i at time $t+1$

is

$$S_0(i,j) = \begin{cases} \text{wait}, & \text{if } |p(t,i)| \le \tau(t, n^{\max}), \\ \text{use active strategy } S(l,i), & \text{otherwise}, \end{cases} \tag{3.39}$$

where the active strategy $S(l,i,j)$ is as follows:

$$S(1,i) = \begin{cases} \text{buy } n_b(t,i,3) \text{ stocks}, & \text{if } p(t,i) \ge p_b(t,i,3), \\ \text{buy } n_b(t,i,2) \text{ stocks}, & \text{if } p(t,i) \ge p_b(t,i,2), \\ & \text{and } p(t,i) < p_b(t,i,3), \\ \text{buy } n_b(t,i,1) \text{ stocks}, & \text{if } p(t,i) \ge p_b(t,i,1) = \tau(t), \\ & \text{and } p(t,i) < p_b(t,i,2), \\ \text{sell } n_s(t,i,1) \text{ stocks}, & \text{if } p(t,i) \le p_s(t,i,1) = -\tau(t), \\ & \text{and } p(t,i) > p_s(t,i,2), \\ \text{sell } n_s(t,i,2) \text{ stocks}, & \text{if } p(t,i) \le p_s(t,i,2), \\ & \text{and } p(t,i) > p_s(t,i,3), \\ \text{sell } n_s(t,i,3) \text{ stocks}, & \text{if } p(t,i) \le p_s(t,i,3), \\ \text{wait}, & \text{otherwise}. \end{cases} \tag{3.40}$$

Here $p(t,i)$ is profitability of investor i at time t defined by (3.2) and the profitability levels are defined by the equalities:

$$p_b(t,i,l) = n_b(t,i,l), p_s(t,i,l) = -n_s(t,i,l), l = 1,2,3. \tag{3.41}$$

In expression (3.40), $n^{\max} = \max(n_b^{\max}, n_s^{\max})$, where $n_b^{\max} = \text{int}(C(t,i)/Z(t))$ and $n_s^{\max} = N(t,i)$.

Expressions (3.41) reflect risk aversion because we accept lesser risk while investing larger assets.

However, we are using the following expression as an alternative:

$$p_b(t,i,l) = \tau(t)l, p_s(t,i,l) = -\tau(t)l, l = 1,2,3. \tag{3.42}$$

Using this strategy, the number of stocks owned by the player i at time $t + 1$ is

$$N(t+1,i)=\begin{cases} N(t,i)+n_b(t,i,3), & \text{if } p(t,i)\geq p_b(t,i,3),\\ N(t)+n_b(t,i,2), & \text{if } p(t,i)\geq p_b(t,i,2),\\ & \text{and } p(t,i)<p_b(t,i,3),\\ N(t)+n_b(t,i,1), & \text{if } p(t,i)\geq p_b(t,i,1)=\tau(t),\\ & \text{and } p(t,i)<p_b(t,i,2),\\ N(t)-n_s(t,i,1), & \text{if } p(t,i)\leq p_s(t,i,1)=-\tau(t),\\ & \text{and } p(t,i)>p_s(t,i,2),\\ N(t)-n_s(t,i,2), & \text{if } p(t,i)\leq p_s(t,i,2),\\ & \text{and } p(t,i)>p_s(t,i,3),\\ N(t)-n_s(t,i,3), & \text{if } p(t,i)\leq p_s(t,i,3),\\ N(t), & \text{if no deal.}\end{cases} \quad (3.43)$$

The buying-selling prices of the player i at time t depends on the buying-selling levels l. Extending single-level conditions (3.8) and (3.9) to the multi-level case of active strategy $S(l,i)$, the buying-selling price levels are as follows:

$$z_b(t,i,l)=\frac{z(t+1,i)}{1-\delta(t)+\alpha(t)+h(t)+p_b(t,i,l)},$$
$$z_s(t,i,l)=\frac{z(t+1,i)}{1-\delta(t)+\alpha(t)+h(t)+p_s(t,i,l)},l=1,2,3. \quad (3.44)$$

Here $0<z_a\leq z_b(t,i,l)\leq z_s(t,i,l)\leq z_b<\infty$.

It follows from (3.30) and (3.31) that:

$$z_b(t,i,l+1)<z_b(t,i,l),$$
$$z_s(t,i,l+1)>z_s(t,i,l), \quad (3.45)$$
$$z_s(t,i,l)>z_b(t,i,l),l=1,2,3.$$

Using (3.2), (3.42) and (3.44) we write buying/selling price levels (3.44) in this form:

$$z_b(t,i,l)=\left(\frac{z(t+1,i)}{1-\delta(t)+\alpha(t)+h(t)+\tau(t)l}\right),$$
$$z_s(t,i,l)=\left(\frac{z(t+1,i)}{1-\delta(t)+\alpha(t)+h(t)-\tau(t)l}\right),l=1,2,3. \quad (3.46)$$

The actual price of a stock at time $t+1$ is defined as a weighted average of the present stock price $Z(t)$ and the price of a previous deal of major stockholders plus the truncated Gaussian random number $\varepsilon(t+1)$ representing the remaining small stockholders. Thus, the actual stock price at time $t+1$ determined by buying-selling actions of a stockholder i is this:

$$Z(t+1,i) = \\ = \begin{cases} (1-a)z_b(t,i,3) + aZ(t) + \varepsilon(t+1), & \text{if } p(t,i) \geq p_b(t,i,3), \\ (1-a)z_b(t,i,2) + aZ(t) + \varepsilon(t+1), & \text{if } p(t,i) \geq p_b(t,i,2), \\ & \text{and } p(t,i) < p_b(t,i,3), \\ (1-a)z_b(t,i,1) + aZ(t) + \varepsilon(t+1), & \text{if } p(t,i) \geq p_b(t,i,1) = \tau(t), \\ & \text{and } p(t,i) < p_b(t,i,2), \\ (1-a)z_s(t,i,1) + aZ(t) + \varepsilon(t+1), & \text{if } p(t,i) \leq p_s(t,i,1) = -\tau(t), \\ & \text{and } p(t,i) > p_s(t,i,2), \\ (1-a)z_s(t,i,2) + aZ(t) + \varepsilon(t+1), & \text{if } p(t,i) \leq p_s(t,i,2), \\ & \text{and } p(t,i) > p_s(t,i,3), \\ (1-a)z_s(t,i,3) + aZ(t) + \varepsilon(t+1) & \text{if } p(t,i) \leq p_s(t,i,3), \\ Z(t) + \varepsilon(t+1) & \text{if no deal.} \end{cases} \quad (3.47)$$

Here a is the coefficient of market inertia.

The deal happens if the selling stockholder has stocks to sell and the buying stockholder has sufficient funds. Expressing conditions (3.47) in terms of buying-selling price levels we write:

$$Z(t+1,i) = \\ = \begin{cases} (1-a)z_b(t,i,3) + aZ(t,i) + \varepsilon(t+1), & \text{if } Z(t,i) \leq z_b(t,i,3), \\ ((1-a)z_b(t,i,2) + aZ(t,i) + \varepsilon(t+1), & \text{if } Z(t,i) \leq z_b(t,i,2), \\ & \text{and } Z(t,i) > z_b(t,i,3), \\ ((1-a)z_b(t,i,1) + aZ(t,i) + \varepsilon(t+1), & \text{if } Z(t) \leq z_b(t,i,1), \\ & \text{and } Z(t,i) > z_b(t,i,2), \\ (1-a)z_s(t,i,1) + aZ(t,i) + \varepsilon(t+1), & \text{if } Z(t,i) \geq z_s(t,i,1), \\ & \text{and } Z(t,i) < z_s(t,i,2), \\ (1-a)z_s(t,i,2) + aZ(t,i) + \varepsilon(t+1), & \text{if } z(t,i) \geq z_s(t,i,2), \\ & \text{and } Z(t,i) < z_s(t,i,3), \\ (1-a)z_s(t,i,3) + aZ(t,i) + \varepsilon(t+1) & \text{if } p(t,i) \geq z_s(t,i,3), \\ Z(t,i) + \varepsilon(t+1) & \text{if no deal.} \end{cases} \quad (3.48)$$

The problem is how to define $Z(t+1)$, if several buying conditions by different stockholders i are satisfied at the same time. A possible solution is to prefer the user which buying price level $z_b(t,i,l)$ is crossed by the actual price line $Z(t)$first assuming that this user has sufficient funds.

It follows from (3.8) that the highest level $z_b(t,i,l=1)$ will be crossed first. Therefore, this condition can be reduced to maximization at the first buying level $l = 1$:

$$i^{\max} = \arg \max_{i} z_b\,(t, i, 1). \tag{3.49}$$

Similar problem is how to define $Z(t+1)$ if several selling conditions by different stockholders i are satisfied at the same time. A solution is to prefer the user which selling price level $z_s(t, i, l)$ is crossed by the actual price line $Z(t)$ first assuming that this user has stocks for sale.

It follows from (3.9) that the lowest level $z_s(t, i, l = 1)$ will be crossed first. Therefore, this condition can be reduced to minimization at the first level $l = 1$:

$$i^{\min} = \arg \min_{i} z_s\,(t, i, 1), \tag{3.50}$$

The actual stock price at time $t + 1$ determined by buying actions of stockholders is defined by this expression:

$$Z_b(t+1) = (1-a) z_b(t, i^{\max}, 1) + aZ(t) + \varepsilon(t+1). \tag{3.51}$$

The actual stock price at time $t + 1$ determined by selling actions of stockholders is

$$Z_s(t+1) = (1-a) z_s(t, i^{\min}, 1) + aZ(t) + \varepsilon(t+1). \tag{3.52}$$

If no buying-selling conditions hold then:

$$Z(t+1) = Z(t) + \varepsilon(t+1). \tag{3.53}$$

The problem remains if both buying and selling conditions are met at the same time. This can happen, since different stockholders are using different prediction rules. Simplest solution would be to set average:

$$Z_a(t+1) = (1-a)\left(Z_b(t+1) + \frac{Z_s(t+1)}{2} + aZ(t) + \epsilon(t+1)\right). \tag{3.54}$$

Then

$$Z(t+1) = \\ = \begin{cases} Z_b(t+1), & \text{if only the buying operation occurs,} \\ Z_s(t+1), & \text{if only the selling operation occurs,} \\ Z_a(t+1), & \text{if both buying and selling operations happen,} \\ Z(t) + \varepsilon(t+1) & \text{if no buying} - \text{selling.} \end{cases} \tag{3.55}$$

Conditions (3.49) and (3.50) reduce the multi-level expression (3.48) to single first level. This is convenient for software testing.

For experimental calculations the average buying-selling levels can be preferred while defining the price $Z(t+1)$ when several buying-selling conditions are satisfied simultaneously. Then the stock price at time $t+1$ determined by buying-selling actions of all stockholders is as follows

$$Z(t+1) = \\ = \begin{cases} \frac{1}{IL(t)} \sum_{i,l \in il(t)} z_b\,(t,i,l) + \varepsilon(t+1), \\ Z(t) + \varepsilon(t+1) & \text{if no buying} - \text{selling}. \end{cases} \tag{3.56}$$

In this expression, the symbol $il(t)$ defines the set of pairs (i,l) which are active at time t according to conditions (3.40) and (3.48). The symbol $IL(t)$ shows the number of elements of the set $il(t)$ defining the number of simultaneous transactions.

However, it is not clear yet if condition (3.56) describes the real stock exchange correctly. Thus, this condition is not implemented yet.

3.3. Trading Rules

In the present version of the PORTFOLIO model, 190 different trading strategies are implemented. These strategies are generated using ten trading rules and nineteen forecasting model. In this chapter, all ten trading rules will be described in detail.

3.3.1. Multi-Stock Operations, Portfolio Problem

In this section, four heuristic trading rules representing personal opinions of some real stockholders with different approaches to risk are described. The advantage is the simplicity of these procedures allowing daily updates. This is important in the short term investing.

Considering longer-term investing, additional trading rules are applied. The first one estimates the risk using bankruptcy probabilities and the utility theory. The second trading rule imitates MPT by maximizing the Sharpe ratio. Advantage of these two trading rules is some theoretical base. The disadvantage is the long computing time. Therefore, in this work, these trading rules are used just for longer term investing. The remaining four trading rules are longer-term extensions of the first four short-term rules.

In the experiments, data is divided in the learning and testing sets. The learning set is for parameter estimation. In the testing set, the price predictions are produced using the parameters defined by the learning set. The length of both sets is about 180 working days each, as usual.

3.3.2. Trading Rule No. 1, Risk-Aware Stockholders: Buying the Best – Selling the Losers by Three Profitability Levels

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

$$j^{\max} = \arg \max_{j} p\,(t, i, j). \quad (3.57)$$

First, the stockholder i sells all nonprofitable stocks:

$$p_s(t, i, j) \leq -\tau(t, i, j), \quad (3.58)$$

and then invests all available funds to buy the most profitable stock. The stockholder i do not sell the stock j, if the expected loss is less than the transaction cost $|p(t, i, j)| < \tau(t, i, j)$. We assume that transaction costs τ are the same for all stocks and do not depend on time. However, extending expression (3.3) of relative transaction costs to multi-stock case we use indexes (t, i, j) instead of (t, n), since these costs depend on the numbers n of stocks j involved in the operation at time t by a stockholder i.

This selling strategy reflects risk-aware users, which keep some less profitable stocks to avoid possible losses if predictions happen to be wrong.

Note that the risk-neutral users sell all the stocks with profitability less then maximal and then invest all available funds in the stock $j^{\max}$, which provides the maximal return. This way they maximize the expected profit. Details are in the next section (3.3.3.).

The investor's i own funds at time t, including the income from selling unprofitable stocks, are expressed as the sum:

$$C_0(t,i) = \sum_j C_0\,(t,i,j), \tag{3.59}$$

where $C_0(t,i,j)$ is defined by this recurrent expression:

$$C_0(t,i,j) = C_0(t-1,i,j) - (N(t,i,j) - N(t-1,i,j))\,Z(t,j). \tag{3.60}$$

The investors' funds available for investing are

$$C(t,i) = C_0(t,i,j) + L(t,i) - B_{\text{sum}}(t,i). \tag{3.61}$$

Here $t = 1,\dots,T$, $L(t,i)$ is the credit limit at time t, and $B_{\text{sum}}(t,i)$ is the borrowed sum defined by multi-stock extension of expression (3.20).

Then we invest all available resources to buy the most profitable stock $j^{\max}$. This means that we sell stocks as the risk aware user but we buy stocks as the risk-neutral one. Thus, the feasible number of stocks $j = j^{\max}$ to buy at time t is as follows:

$$n_b(t,i,j^{\max}) = \text{int}\left(\frac{C(t,i)}{Z(t,j^{\max})}\right), \text{if } p(t,i,j^{\max}) > \tau(t,i,j^{\max}). \tag{3.62}$$

The general buying-selling strategy $S_0(i,j)$ of the investor i at time $t+1$ is

$$S_0(i,j) = \begin{cases} \text{wait}, & \text{if } p(t,i,j^{\max}) \le \tau(t,i,j^{\max}), \\ \text{use active strategy } S(i,j), & \text{otherwise}, \end{cases} \tag{3.63}$$

where the active strategy $S(l,i,j)$ is as follows:

$$S(i,j) = \\ = \begin{cases} \text{sell } n_s(t,i,j,1), \text{ stocks,} & \text{if } p(t,i,j) \le p_s(t,i,j,1) = -\tau(t,j), \\ & \text{and } p(t,i,j) > p_s(t,i,j,2), \\ \text{sell } n_s(t,i,j,2) \text{ stocks,} & \text{if } p(t,i,j) \le p_s(t,i,j,2), \\ & \text{and } p(t,i,j) > p_s(t,i,j,3), \\ \text{sell } n_s(t,i,j,3) \text{ stocks,} & \text{if } p(t,i,j) \le p_s(t,i,j,3), \\ \text{buy } n_b(t,i,j^{\max}) \text{ stocks by all funds.} & \end{cases} \quad (3.64)$$

Here $p(t,i,j)$ is profitability of stock j of investor i at time t defined by multi-stock extension of (3.2) and the profitability levels are defined by these expressions:

$$p_s(t,i,j,l) = -\tau(ij)\, l, l = 1,2,3. \quad (3.65)$$

Using the strategy $S_0(i,j)$, the number of stocks j owned by the player i at time $t+1$ is

$$N(t+1,i,j) = \\ = \begin{cases} N(t,i,j) + n_b(t,i,j^{\max}), & \text{if } p(t,i,j^{\max}) > \tau(t,i,j^{\max}), \\ N(t,j) - n_s(t,i,j,1), & \text{if } p(t,i,j) \le p_s(t,i,j,1) = -\tau(t,j), \\ & \text{and } p(t,i,j) > p_s(t,i,j,2), \\ N(t,j) - n_s(t,i,j,2), & \text{if } p(t,j,i) \le p_s(t,i,j,2), \\ & \text{and } p(t,i,j) > p_s(t,i,j,3), \\ N(t,j) - n_s(t,i,j,3), & \text{if } p(t,i,j) \le p_s(t,i,j,3), \\ N(t,j), & \text{if no deal.} \end{cases} \quad (3.66)$$

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

The buying-selling prices of stock j of the player i at time t depends on the buying-selling levels l. Using (3.2), (3.44) and (3.46) we write buying/selling price levels in this form:

$$z_b(t,i,j^{\max}) = \frac{z(t+1,i,j^{\max})}{1-\delta(t)+\alpha(t)+h(t)+\tau(t,j^{\max})}, \\ z_s(t,i,j,l) = \frac{z(t+1,i,j)}{1-\delta(t)+\alpha(t)+h(t)-\tau(t,j)l}, l = 1,2,3. \quad (3.67)$$

The actual price of a stock at time $t+1$ is defined as the price of a previous deal of major stockholders plus the truncated Gaussian noise representing the remaining small stockholders. Thus, the actual stock j price at time $t+1$ determined by buying-selling actions of a stockholder i is this:

$$Z(t+1,i,j) = \\ = \begin{cases} (1-a)z_b(t,i,j^{\max}) + aZ(t,j^{\max}) + \varepsilon(t+1), & \text{if } p(t,i,j^{\max}) > 0, \\ (1-a)z_s(t,i,j,1) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } p(t,i,j) \le p_s(t,i,j,1) = -\tau(t,j), \\ & \text{and } p(t,i,j) > p_s(t,i,j,2), \\ (1-a)z_s(t,i,j,2) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } p(t,i,j) \le p_s(t,i,j.2), \\ & \text{and } p(t,i.j) > p_s(t,i,j,3), \\ (1-a)z_s(t,i,j,3) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } p(t,i,j) \le p_s(t,i,j,3), \\ Z(t,j) + \varepsilon(t+1,j), & \text{if no deal.} \end{cases} \quad (3.68)$$

Expressing conditions (3.68) in terms of buying-selling price levels we write:

$$Z(t+1,i,j) = \\ = \begin{cases} (1-a)z_b(t,i,j^{\max}) + aZ(t,j^{\max}) + \varepsilon(t+1), & \text{if } Z(t,j^{\max}) \le z_b(t,i,j^{\max}), \\ (1-a)z_s(t,i,j,1) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } Z(t,i,j) \ge z_s(t,i,j,1), \\ & \text{and } Z(t,i,j) < z_s(t,i,j,2), \\ (1-a)z_s(t,i,j,2) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } Z(t,i,j) \ge z_s(t,i,j,2), \\ & \text{and } Z(t,i,j) < z_s(t,i,j,3), \\ (1-a)z_s(t,i,j,3) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } Z(t,i,j) \ge z_s(t,i,j,3), \\ Z(t,j) + \varepsilon(t+1,j), & \text{if no deal.} \end{cases} \quad (3.69)$$

Here the noise $\varepsilon(t+1,j)$ is generated by the truncated Gaussian distribution with minimal values restricted by the following multi-stock version:

$$Z(t+1,i,j) \ge \rho_j + 3\tau_0, \quad (3.70)$$

where $\rho_j > 0$ is the minimal stock price "insolvency level". This inequality is implemented by ignoring those ϵ values, which are crossing the lower limit (3.70).

The problem is how to define $Z(t+1,j)$, if several buying conditions by different stockholders i are satisfied at the same time. A possible solution is to prefer the user which buying price level $z_b(t,i,j^{\max})$ is crossed by the actual price line $Z(t,j^{\max})$ first.

Therefore, this condition can be reduced to this maximization:

$$i^{\max} = \arg\max_i z_b\,(t, i, j^{\max}). \tag{3.71}$$

Similar problem is how to define $Z(t+1)$ if several selling conditions by different stockholders i are satisfied at the same time. A solution is to prefer the user which selling price level $z_s(t,i,j,l)$ is crossed by the actual price line $Z(t,j)$ first.

It follows from (3.9) that the lowest level $z_s(t,i,j,l=1)$ will be crossed first. Therefore, this condition can be reduced to minimization at the first level $l=1$:

$$i^{\min} = \arg\min_i z_s\,(t, i, j^{\min}, 1), \tag{3.72}$$

The actual stock price at time $t+1$ determined by buying actions of stockholders is defined by this expression:

$$Z_b(t+1, j^{\max}) = (1-a)z_b(t, i^{\max}, j^{\max}) + aZ(t,j) + \varepsilon(t+1). \tag{3.73}$$

The actual stock price at time $t+1$ determined by selling actions of stockholders is

$$Z_s(t+1, j^{\min}) = (1-a)z_s(t, i^{\min}, j^{\min}, l) + aZ(t,j) + \varepsilon(t+1,j). \tag{3.74}$$

If no deal then:

$$Z(t+1,j) = Z(t,j) + \varepsilon(t+1,j). \tag{3.75}$$

Suppose that for some stock j^{both} both buying and selling conditions are met at the same time. This can happen, since different stockholders are using different prediction rules. Simplest solution would be to set average:

$$Z_a(t+1, j^{\text{both}}) = \frac{(1-a)\left(Z_b(t+1, j^{\text{both}}) + Z_s(t+1, j^{\text{both}})\right)}{2} + \tag{3.76}$$
$$+ aZ(t, j^{\text{both}}) + \varepsilon(t+1, j^{\text{both}}).$$

Then:

$$Z(t+1,j) =$$
$$= \begin{cases} Z_b(t+1, j^{\max}), & \text{if only the buying operation occurs,} \\ Z_s(t+1, j^{\min}), & \text{if only the selling operation occurs,} \\ Z_a(t+1, j^{\text{both}}), & \text{if both buying and selling operations happen,} \\ Z(t,j) + \varepsilon(t+1,j), & \text{if no deal.} \end{cases} \tag{3.77}$$

3.3.3. Trading Rule No. 2, Risk-Aware Stockholders: Buying the Best – Selling All the Losers

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

$$j^{\max} = \arg\max_j p\,(t, i, j). \tag{3.78}$$

First, the stockholder i sells all nonprofitable stocks:

$$p_s(t, i, j) \leq -\tau(t, i, j), \tag{3.79}$$

and then invests all available funds to buy the most profitable stock. The stockholder i do not sell the stock j, if the expected loss is less than the transaction cost $|p(t, i, j)| < \tau(t, i, j)$. We assume that transaction costs τ are the same for all stocks and do not depend on time. However, extending expression (3.3) of relative transaction costs to multi-stock case we use indexes (t, i, j) instead of (t, n), since these costs depend on the numbers n of stocks j involved in the operation at time t by a stockholder i.

This selling strategy reflects risk-aware users, which keep some less profitable stocks to avoid possible losses if predictions happen to be wrong.

Note that the risk-neutral users sell all the stocks with profitability less then maximal and then invest all available funds in the stock $j^{\max}$, which provides the maximal return. This way they maximize the expected profit. Details are in the next section (3.3.4.).

The investors' i own funds at time t, including the income from selling unprofitable stocks, are expressed as the sum:

$$C_0(t, i) = \sum_j C_0\,(t, i, j), \tag{3.80}$$

where $C_0(t, i, j)$ is defined by this recurrent expression:

$$C_0(t, i, j) = C_0(t-1, i, j) - (N(t, i, j) - N(t-1, i, j))\,Z(t, j). \tag{3.81}$$

The investors' funds available for investing are

$$C(t,i) = C_0(t,i,j) + L(t,i) - B_{\text{sum}}(t,i). \tag{3.82}$$

Here $t = 1,\ldots,T$, $L(t,i)$ is the credit limit at time t, and $B_{\text{sum}}(t,i)$ is the borrowed sum defined by multi-stock extension of expression (3.20).

Thus, the feasible number of stocks $j = j^{\max}$ to buy at time t is as follows:

$$n_b(t,i,j^{\max}) = \text{int}\left(\frac{C(t,i)}{Z(t,j^{\max})}\right), \text{ if } p(t,i,j^{\max}) > \tau(t,i,j^{\max}). \tag{3.83}$$

The general buying-selling strategy $S_0(i,j)$ of the investor i at time $t+1$ is

$$S_0(i,j) = \begin{cases} \text{wait}, & \text{if } p(t,i,j^{\max}) \le \tau(t,i,j^{\max}), \\ \text{use active strategy } S(i,j), & \text{otherwise}, \end{cases} \tag{3.84}$$

where the active strategy $S(l,i,j)$ is as follows:

$$S(i,j) =$$
$$= \begin{cases} \text{sell } N(t,i,j), \text{ stocks}, & \text{if } p(t,i,j) \le -\tau(t,j), \\ \text{buy } n_b(t,i,j^{\max}) \text{ stocks by all funds}. \end{cases} \tag{3.85}$$

Here $p(t,i,j)$ is profitability of stock j of investor i at time t defined by multi-stock extension of (3.2). Using the strategy $S_0(i,j)$, the number of stocks j owned by the player i at time $t+1$ is

$$N(t+1,i,j) =$$
$$= \begin{cases} N(t,i,j) + n_b(t,i,j^{\max}), & \text{if } p(t,i,j^{\max}) > \tau(t,i,j^{\max}) \\ 0, & \text{if } p(t,i,j) \le -\tau(t,j) \\ & \text{and } p(t,i,j^{\max}) \le \tau(t,i,j^{\max}) \\ N(t,j), & \text{if no deal.} \end{cases} \tag{3.86}$$

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

The buying-selling prices of stock j of the player i at time t depends on the buying-selling levels l. Using (3.2), (3.44) and (3.46) we write buying/selling price levels in this form:

$$z_b(t,i,j^{\max}) = \frac{z(t+1,i,j^{\max})}{1-\delta(t)+\alpha(t)+h(t)+\tau(t,j^{\max})},$$
$$z_s(t,i,j) = \frac{z(t+1,i,j)}{1-\delta(t)+\alpha(t)+h(t)-\tau(t,j)}, l = 1,2,3. \quad (3.87)$$

The actual price of a stock at time $t+1$ is defined as the price of a previous deal of major stockholders plus the truncated Gaussian noise representing the remaining small stockholders. Thus, the actual stock j price at time $t+1$ determined by buying-selling actions of a stockholder i is this:

$$Z(t+1,i,j) =$$
$$= \begin{cases} (1-a)z_b(t,i,j^{\max}) + aZ(t,j^{\max}) + \varepsilon(t+1), & \text{if } p(t,i,j^{\max}) > 0, \\ (1-a)z_s(t,i,j) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } p(t,i,j) \le p(t,i,j) \le -\tau(t,j), \\ Z(t,j) + \varepsilon(t+1,j) & \text{if no deal.} \end{cases} \quad (3.88)$$

Expressing conditions (88) in terms of buying-selling price levels we write:

$$Z(t+1,j,i) =$$
$$= \begin{cases} (1-a)z_b(t,i,j^{\max}) + aZ(t,j^{\max}) + \varepsilon(t+1), & \text{if } Z(t,j^{\max}) \le z_b(t,i,j^{\max}), \\ (1-a)z_s(t,i,j) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } Z(t,i,j) \ge z_s(t,i,j), \\ Z(t,j) + \varepsilon(t+1,j), & \text{if no deal.} \end{cases} \quad (3.89)$$

The noise $\varepsilon(t+1,j)$ is generated by the truncated Gaussian distribution with minimal values restricted by the following multi-stock version of condition (3.11). This inequality is implemented by ignoring those ε values which are crossing the lower limit (3.70).

The actual stock price at time $t+1$ determined by buying actions of stockholders is defined by this expression:

$$Z_b(t+1,j^{\max}) = (1-a)z_b(t,i^{\max},j^{\max}) + aZ(t,j) + \varepsilon(t+1). \quad (3.90)$$

The actual stock price at time $t+1$ determined by selling actions of stockholders is

$$Z_s(t+1,j^{\min}) = (1-a)z_s(t,i^{\min},j^{\min},l) + aZ(t,j)\varepsilon(t+1,j). \quad (3.91)$$

If no deal then:

$$Z(t+1,j) = Z(t,j) + \varepsilon(t+1,j). \quad (3.92)$$

Suppose that for some stock j^{both} both buying and selling conditions are met at the same time. This can happen, since different stockholders are using different prediction rules. Simplest solution would be to set average

$$Z_a(t+1, j^{\text{both}}) = \\ = \frac{(1-a)\left(Z_b(t+1, j^{\text{both}}) + Z_s(t+1, j^{\text{both}})\right)}{2} \\ + aZ(t, j^{\text{both}}) + \varepsilon(t+1, j^{\text{both}}). \tag{3.93}$$

Then:

$$Z(t+1, j) = \\ = \begin{cases} Z_b(t+1, j^{\max}), & \text{if only the buying operation occurs,} \\ Z_s(t+1, j^{\min}), & \text{if only the selling operation occurs,} \\ Z_a(t+1, j^{\text{both}}), & \text{if both buying and selling operations happen,} \\ Z(t, j) + \varepsilon(t+1, j) & \text{if no deal.} \end{cases} \tag{3.96}$$

3.3.4. Trading Rule No. 3, Risk-Neutral Stockholders: Buying the Best – Selling All the Rest

The risk-neutral stockholders use all available resources to buy stock $j^{\max}$, which provides the highest expected profit:

$$j^{\max} = \arg\max_j p\,(t, i, j). \tag{3.97}$$

Denote by $J(\tau)$ a subset of stocks with profitability less or equal to the best minus the relative transaction cost:

$$J(\tau) = \{j : p(t.i.j) \le p(t, i, j^{\max}) - \tau(t, n_s(t, i, j))\}, \tag{3.98}$$

where $n_s(t, i, j)$ is the number of stocks j for sale at time t by stockholder i. Here, defining the relative transaction cost, we use the longer symbol $\tau(t, n_s(t, i, j))$ instead of the shorter one $\tau(t, i, j)$ to show the number of stocks $n_s(t, i, j)$ directly.

First, the risk-neutral stockholder is selling the stocks $j \in J(\tau)$ to raise funds for buying the single most profitable stock $j^{\max}$.

Stockholders do nothing, if the maximal expected profit is less than the transaction cost $C(t, i)p(t, i, j^{\max}) < \tau(t, n_b(t, j^{max}))$ and do not sell if the maximal expected losses are less than $C(t, i)p(t, i, j^{\min}) < \tau(t, n_s(t, j^{\min})$.

Thus, the number of stocks $j = j^{\max}$ to buy at time t is as follows:

$$n_b(t, j^{\max}) = \operatorname{int}\left(\frac{C(t,i)}{Z(t, j^{\max})}\right),$$
$$\text{if } p(t,i,j^{\max}) \geq \tau(t, n_b(t, j^{\max})). \tag{3.99}$$

We do not sell, if the maximal expected losses are less than the transaction cost $N(t,i,j^{\min})Z(t,i)p(t,i,j^{\min}) < -\tau(t, n_s(t, j^{\min}))$, where $N(t,i,j)$ is the number of stocks j available at time t. The feasible number of stocks j to sell at time t is

$$n_s(t,i,j^{\min}) = N(t,i,j), \text{if } p(t,i,j) \leq p_s(t,i,j). \tag{3.100}$$

The general buying-selling strategy $S_0(i,j)$ of the investor i at time $t+1$ is

$$S_0(i,j) = $$
$$= \begin{cases} \text{wait}, & \text{if } p(t,i,j^{\max}) \leq \tau(t,i,j^{\max}), \\ \text{use active strategy } S(i,j), & \text{otherwise}, \end{cases} \tag{3.101}$$

where the active strategy $S(l,i,j)$ is as follows:

$$S(1,i,j) = \begin{cases} \text{sell } n_s(t,i,j) \text{ stocks}, & \text{if } j \in J(\tau), \\ \text{buy } n_b(t,i,j^{\max}) \text{ stocks}, & \text{by all funds}. \end{cases} \tag{3.102}$$

Here $p(t,i,j)$ is profitability of stock j of investor i at time t defined by multi-stock extension of (3.2).

Using these strategies, the number of stocks j owned by the player i at time $t+1$ is

$$N(t+1,i,j) =$$
$$= \begin{cases} N(t,i,j) + n_b(t,i,j^{\max}), & \text{if if } p(t,i,j^{\max}) \geq p_b(t,i,j^{\max}), \\ N(t,j) - n_s(t,i,j), & \text{if } j \in J(\tau), (10) \\ N(t,j), & \text{if no deal}. \end{cases} \tag{3.103}$$

The buying-selling price levels we define by this expression:

$$z_b(t,i,j^{\max}) = \frac{z(t+1,i,j^{\max})}{1 - \delta(t) + \alpha(t) + h(t) + p_b(t,i,j^{\max})},$$
$$z_s(t,i,j) = \frac{z(t+1,i,j)}{1 - \delta(t) + \alpha(t) + h(t) + p_s(t,i,j)}, j \in J(\tau). \tag{3.104}$$

The actual price of a stock at time $t+1$ is defined as the price of a previous deal of major stockholders plus the truncated Gaussian noise representing the remaining small stockholders. Thus, the actual stock j price at time $t+1$ determined by buying-selling actions of a stockholder i is this:

$$Z(t+1,i,j) = \\ = \begin{cases} (1-a)z_b(t,i,j^{\max}) + aZ(t,j^{\max}) + \varepsilon(t+1), & \text{if } p(t,i,j^{\max}) \geq p_b(t,i,j^{\max}), \\ (1-a)z_s(t,i,j) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } j \in J(\tau), \\ Z(t,j) + \varepsilon\varepsilon(t+1,j) & \text{if no deal.} \end{cases} \quad (3.105)$$

Expressing conditions (3.68) in terms of buying-selling price levels we write:

$$Z(t+1,i,j) = \\ = \begin{cases} (1-a)z_b(t,i,j^{\max}) + aZ(t,j^{\max}) + \varepsilon(t+1), & \text{if } Z(t,j) \leq z_b(t,i,j^{\max}), \\ (1-a)z_s(t,i,j) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } j \in J(\tau), \\ Z(t,j) + \varepsilon(t+1,j) & \text{if no deal.} \end{cases} \quad (3.106)$$

The problem is how to define $Z(t+1,j)$, if several buying conditions by different stockholders i are satisfied at the same time. A possible solution is to prefer the user which buying price level $z_b(t,i,j)$ is crossed by the actual price line $Z(t,j)$ first.

It follows from (3.8) that the highest level $z_b(t,i,j^{\max})$ will be crossed first. Therefore, this condition can be reduced to maximization:

$$i^{\max} = \arg\max_i z_b\,(t,i,j^{\max}). \quad (3.107)$$

Similar problem is how to define $Z(t+1)$ if several selling conditions by different stockholders i are satisfied at the same time. A solution is to prefer the user which selling price level $z_s(t,i,j)$ is crossed by the actual price line $Z(t,j)$first.

It follows from (3.9) that the lowest level $z_s(t,i,j)$ will be crossed first. Therefore, this condition can be reduced to minimization:

$$i^{\min} = \arg\min_i z_s\,(t,i,j^{\min}). \quad (3.108)$$

The actual stock price at time $t+1$ determined by buying actions of stockholders is defined by this expression:

$$Z_b(t+1, j^{\max}) = (1-a)z_b(t, i^{\max}, j^{\max}) + aZ(t,j) + \varepsilon(t+1). \quad (3.109)$$

The actual stock price at time $t+1$ determined by selling actions of stockholders is

$$Z_s(t+1, j) = (1-a)z_s(t, i^{\min}, j) + aZ(t,j)\varepsilon(t+1,j), \quad (3.110)$$

where $j \in J(\tau)$.

If no buying-selling conditions hold, then:

$$Z(t+1, j) = Z(t,j) + \varepsilon(t+1,j). \quad (3.111)$$

The problem remains if for some stock j^{both} both buying and selling conditions are met at the same time. This can happen, since different stockholders are using different prediction rules. Simplest solution would be to set average:

$$\begin{aligned} Z_a(t+1, j^{\text{both}}) &= \\ &= \frac{(1-a)\left(Z_b(t+1, j^{\text{both}}) + Z_s(t+1, j^{\text{both}})\right)}{2} \\ &+ aZ(t) + \varepsilon(t+1,j). \end{aligned} \quad (3.112)$$

Then:

$$Z(t+1,j) = \\ = \begin{cases} Z_b(t+1, j^{\max}), & \text{if the buying operation occurs,} \\ Z_s(t+1, j), j \in J(\tau) & \text{if the selling operation occurs,} \\ Z_a(t+1, j^{\text{both}}), & \text{if both buying and selling operations happen,} \\ Z(t,j) + \varepsilon(t+1,j) & \text{if no buying} - \text{selling of stock } j. \end{cases} \quad (3.113)$$

All these operations are controlled by the general buying-selling strategy $S_0(t,j)$.

3.3.5. Trading Rule No. 4, Risk-Averse Stockholders: Selling and Buying in Proportion to Profitability

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

$$j_{+}^{\max} = \arg \max_{j \in J_{+}} p\,(t, i, j), \tag{3.114}$$

and

$$j_{-}^{\min} = \arg \min_{j \in J_{-}} p\,(t, i, j). \tag{3.115}$$

First, we sell stocks in proportion to $l = 1, \ldots, j_{-}^{\min}$ selling profitability levels $p_s(t, i, l) = p(t, i, j = l), l = 1, \ldots, j_{-}^{\min}$. Then we use all accumulated resources to buy stocks in proportion to $l = 1, \ldots, j_{+}^{\max}$ profitability levels $p_b(t, i, l) = p(t, i, j = l), l = 1, \ldots, j_{-}^{\max}$.

The investors' i own funds at time t, including the income from stocks sold at time t, are expressed as the sum:

$$C_0(t, i) = \sum_{j} C_0\,(t, i, j), \tag{3.116}$$

where $C_0(t, i, j)$ is defined by this recurrent expression.

$$C_0(t, i, j) = C_0(t-1, i, j) - (N(t, i, j) - N(t-1, i, j))\,Z(t, j). \tag{3.117}$$

The investors' funds available for investing are

$$C(t, i) = C_0(t, i, j) + L(t, i) - B_{\text{sum}}(t, i), \tag{3.118}$$

here $t = 1, \ldots, T$, $L(t, i)$ is the credit limit at time t, and $B_{\text{sum}}(t, i)$ is the borrowed sum defined by multi-stock extension of expression (3.20). This enables us to distribute all available resources in proportion to the profitability of stocks.

For example, at selling level l we sell:

$$n_s(t, i, l) = \text{int}(N(t, i, j)\frac{2l}{J_s(J_s+1)}) \tag{3.119}$$

of stocks, and at buying level l we buy:

$$n_b(t, i, l) = \text{int}(C(t, i)\frac{2l}{J_b(J_b+1)Z(t, i)}) \tag{3.120}$$

of stocks using a part $2l/J_b(J_b+1)$ of available resources. We apply the standard rounding up procedure for number of stocks n to sell and buy. The balance is corrected at the first level $l = 1$.

Suppose that transaction costs τ are the same for all stocks and do not depend on time. However, extending expression (3.3) of relative transaction costs to multi-stock case we use indexes (t, i, j) instead of (t, n), since these costs depend on the numbers n of stocks j involved in the operation at time t by a stockholder i.

We do not sell/buy the stock j, if the expected loss/profit is less than the transaction cost $|p(t, i, j)| < \tau(t, i, j)$.

The general buying-selling strategy $S_0(i, j)$ of the investor i at time $t + 1$ is different to that described in the single-stock section *3.2.7. Multi-level operations* because here investors need some additional rules how to distribute limited resources between different stocks.

$$S_0(i,j) =$$
$$= \begin{cases} \text{wait}, & \text{if } |p(t,i,j)| \le \tau(t,i,j) \text{ for all } j, \\ \text{use active strategy } S(i,j), & \text{otherwise}, \end{cases} \quad (3.121)$$

where the active strategy $S(l, i, j)$ is as follows:

$$S(i,j) =$$
$$= \begin{cases} \text{sell } n_s(t,i,l), \text{ stocks}, & \text{if } p(t,i,j) = p_s(t,i,l)\ l = 1, \dots, j_-^{\min}, \\ \text{buy } n_b(t,i,l), \text{ stocks}, & \text{if } p(t,i,j) = p_b(t,i,l), l = 1, \dots, j_-^{\max}. \end{cases} \quad (3.122)$$

Using the strategy $S_0(i, j)$, the number of stocks j owned by the player i at time $t + 1$ is

$$N(t+1,i,j) =$$
$$= \begin{cases} N(t,j) - n_s(t,i,1), & \text{if } p(t,i,j) = p_s(t,i,1), l = 1, \dots, j_-^{\min}, \\ N(t,j) + n_b(t,i,l), & \text{if } p(t,i,j) = p_b(t,i,l), l = 1, \dots, j_-^{\max}, \\ N(t,j), & \text{if no deal.} \end{cases} \quad (3.123)$$

The buying-selling prices of stock j of the player i at time t depends on the buying-selling levels l. Using (3.2), (3.44) and (3.46) we write buying/selling price levels in this form:

$$z_b(t,i,l) = \frac{z(t+1,i,j)}{1 - \delta(t) + \alpha(t) + h(t) + p_b(t,i,l)}, l = 1, \dots, j_-^{\max}, \quad (3.124)$$

$$z_s(t,i,l) = \frac{z(t+1,i,j)}{1-\delta(t)+\alpha(t)+h(t)-p_s(t,i,l)}, l = 1,\dots,j_-^{\min}.$$

The actual price of a stock at time $t+1$ is defined as the price of a previous deal of major stockholders plus the truncated Gaussian noise representing the remaining small stockholders. Thus, the actual stock j price at time $t+1$ determined by buying-selling actions of a stockholder i is this

$$Z(t+1,i,j) = \\ = \begin{cases} (1-a)z_s(t,i,j,1) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } p(t,i,j) = p_s(t,i,1), l = 1,\dots,j_-^{\min}, \\ (1-a)z_b(t,i,j,l) + aZ(t,j) + \varepsilon(t+1,j), & \text{if } p(t,i,j) = p_b(t,i,l)\ l = 1,\dots,j_-^{\max}, \\ Z(t,j) + \varepsilon(t+1,j), & \text{if no deal.} \end{cases} \tag{3.125}$$

Expressing conditions (3.68) in terms of buying-selling price levels we write:

$$Z(t+1,j,i) = \\ = \begin{cases} (1-a)z_s(t,i,l) + aZ(t,l) + \varepsilon(t+1,l), & \text{if } Z(t,l) \geq z_s(t,i,l)\ l = 1,\dots,j_-^{\min}, \\ (1-a)z_b(t,i,l) + aZ(t,l) + \varepsilon(t+1,l), & \text{if } Z(t,l) \leq z_b(t,i,l)\ l = 1,\dots,j_-^{\max}, \\ Z(t,j) + \varepsilon(t+1,j), & \text{if no deal.} \end{cases} \tag{3.126}$$

The problem is how to define $Z(t+1)$ if several selling conditions by different stockholders i are satisfied at the same time. A solution is to prefer the user which selling price level $z_s(t,i,l)$ is crossed by the actual price line $Z(t,l)$ first.

It follows from (3.9) that the lowest level $z_s(t,i,l=1)$ will be crossed first. Therefore, this condition can be reduced to minimization at the first level $l = 1$:

$$i^{\min} = \arg \min_i z_s\,(t,i,1=1). \tag{3.127}$$

The actual stock price at time $t+1$ determined by buying actions of stockholders is defined by this expression:

$$Z_b(t+1,l) = \\ = (1-a)z_b(t,i,l) + aZ(t,l) + \varepsilon(t+1)l = 1,\dots,j_-^{\max}. \tag{3.128}$$

The actual stock price at time $t+1$ determined by selling actions of stockholders is

$$Z_s(t+1,l) = \tag{3.129}$$

$$= (1-a)z_s(t,i,l) + aZ(t,l) + \varepsilon(t+1,l)\ l = 1,\dots,j_{-}^{\min}.$$

If no deal then:

$$Z(t+1,j) = Z(t,j) + \varepsilon(t+1,j) \text{ for all } j. \tag{3.130}$$

Suppose that for some stock j^{both} both buying and selling conditions are met at the same time. This can happen, since different stockholders are using different prediction rules. Simplest solution would be to set average.

3.4. Longer-Term Investment

In the previous sections, we regarded short term investing by daily decisions. The traditional portfolio problem considers optimal longer-term diversity by defining optimal sharing of available resources between different assets. This can be performed using the individual utility functions, too. This utility function approach is discussed in the next two sections. In the third section, we shall consider the same problem by maximizing the Sharp ratio, following the MPT. Note that in this section different symbols are used since we regard different problems.

The idea of longer term investment to define model parameters by some learning set, for example three, six or twelve months and then use the model defining the future investment strategies. In this work we estimate the goodness of different longer term strategies using the test set of the same duration. In addition, to these specific longer term strategies we apply this longer term approach to all four short term strategies.

3.4.1. Trading Rule No. 5, Individual Approach: Defining Risk by Survival Probabilities and Individual Utility Function

An important part of optimal investment is the definition of individual utility functions that determine particular investors' profit-to-risk relation (Fishburn, 1964). Here we consider an illustrative example how to invest some fixed capital in Certificates of Deposit (CD) and Stocks.

The portfolio problem is to maximize the average utility of wealth. That is obtained by optimal distribution of available capital between different objects with uncertain parameters (Mockus et al., 1997). Denote by x_i the part of the capital invested into an object i. The returned wealth is

$$y_i = c_i x_i.$$

Here

$$c_i = 1 + \delta(t) + \beta(t, i),$$

and

$$\beta(t, i) = \frac{Z_i(t) - Z_i(t-1)}{Z_i(t)}, \tag{3.131}$$

where the $\beta(t, i)$ is the relative stock i price change at time t.

Denote by $p_i = 1 - q_i$ the reliability of investment. Here q_i is the insolvency probability. $u(y)$ is the utility the wealth y. Denote by $U(x)$ the expected utility function. $U(x)$ depends on the capital distribution $x = (x_i, \dots, x_n), \sum_i = 1, x_i \geq 0$. If the wealth is discrete $y = y^k, k = 1, \dots, M$, the expected utility function:

$$U(x) = \sum_{k=1}^{M} u\,(y^k) p(y^k). \tag{3.132}$$

Here M is the number of discrete values of wealth $y^k p_x(y^k)$ is the probability that the wealth y^k will be returned, if the capital distribution is x. We search for such capital distribution x which provides the greatest expected utility of the returned wealth:

$$\max_x U\,(x), \tag{3.133}$$

$$\sum_{i=1}^{n} x_i = 1, x_i \geq 0. \tag{3.134}$$

3.4.1.1 Investment in CD

One may define probabilities $p(y^j)$ of discrete values of wealth $y^j, j = 1,2, \dots$ by exact expressions. For example:

$$p(y^0) = \prod_i q_i,$$

$$p(y^1) = p_1 \prod_{i \neq 1} q_i,$$

$$p(y^2) = p_2 \prod_{i \neq 2} q_i,$$

$$\ldots\ldots\ldots\ldots \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$p(y^n) = p_n \prod_{i \neq n} q_i, \tag{3.135}$$

$$p(y^{n+1}) = p_1 p_2 \prod_{i \neq 1, i \neq 2} q_i,$$

$$p(y^{n+2}) = p_1 p_3 \prod_{i \neq 1, i \neq 3} q_i,$$

$$\ldots\ldots\ldots\ldots \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Here $y^0 = 0$, $y^1 = a_1 x_1$, $y^2 = a_2 x_2$, $y^n = a_n x_n$, $y^{n+1} = a_1 x_1 + a_2 x_2$, $y^{n+2} = a_1 x_1 + a_3 x_3$. From expression (3.135)

$$U(x) = \sum_{k=1}^{M} u\,(y^k) p(y^k). \tag{3.136}$$

Here M is the number of different values of wealth y.

3.4.1.2 Investment in CD and stocks

Investing in CD, the interests α_i are defined by contracts. Only the reliability $p_i, i = 1, \ldots, n$ of banks are uncertain. Investing in stocks, in addition to reliability $p_i, i = n + j, j = 1, \ldots, m$ of companies, their future stock rates are uncertain, too. The predicted stock rates are defined by a coefficient a_i that shows the relation between the present and the 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 $a_i^l, l = 1, \ldots, L$ with corresponding estimated probabilities

$p_i^l, \sum_{l=1}^{L} p_i^l = 1, p_i^l \geq 0.$

In this case, one may define probabilities $p(y^i)$ of discrete values of wealth $y^i, i = 1, \dots, n+m$ by exact expressions. The expressions for CD remain the same. Therefore, we shall consider only stocks assuming that $n = 0$ and $L = 2$. Then:

$$p(y^0) = \prod_i q_i,$$

$$p(y^1) = p_1 p_1^1 \prod_{i \neq 1} q_i,$$

$$p(y^2) = p_1 p_1^2 \prod_{i \neq 1} q_i,$$

$$p(y^3) = p_2 p_2^1 \prod_{i \neq 2} q_i,$$

$$p(y^4) = p_2 p_2^2 \prod_{i \neq 2} q_i, \tag{3.137}$$

$$\dots\dots\dots \quad \dots\dots\dots\dots\dots\dots,$$

$$p(y^{2n-1}) = p_n p_n^1 \prod_{i \neq n} q_i,$$

$$p(y^{2n}) = p_n p_n^2 \prod_{i \neq n} q_i,$$

$$p(y^{2n+1}) = p_1 p_1^1 p_2 p_2^1 \prod_{i \neq 1, i \neq 2} q_i,$$

$$p(y^{2n+2}) = p_1 p_1^2 p_2^p 2^2 \prod_{i \neq 1, i \neq 2} q_i,$$

$$\dots\dots\dots\dots\dots \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots$$

Here $y^0 = 0$, $y^1 = a_1^1 x_1$, $y^2 = a_1^2 x_1$, $y^3 = a_2^1 x_2$, $y^4 = a_2^2 x_2$, $y^{2n-1} = a_n^1 x_n$, $y^{2n} = a_n^2 x_n$, $y^{2n+1} = a_1^1 x_1 + a_2^1 x_2$, $y^{2n+2} = a_1^2 x_1 + a_2^2 x_2$. The reliability p_i, the stock rate predictions a_i^l and their estimated probabilities p_i^l are defined by experts, possibly, with the help of time series models such as ARMA. For example, maximal values of multi-step prediction are considered as "optimistic" estimates and the minimal values-as "pessimistic" ones. The average values of multi-step prediction are regarded as "realistic" estimates.

Here is a simplest illustration were $n = m = 1$ and $L = 2$. In this case from (3.135) (3.137) the probabilities $p(y^k)$ of wealth returns y^k, $k = 0, \ldots, 5$ are

$$p(y^0) = q_1 q_2,$$
$$p(y^1) = p_1 q_2,$$
$$p(y^2) = p_2 p_2^1 q_1,$$
$$p(y^3) = p_2 p_2^2 q_1,$$
$$p(y^4) = p_2 p_2^1 p_1,$$
$$p(y^5) = p_2 p_2^2 p_1.$$

Here $y^0 = 0$, $y_1 = a_1 x_1$, $y^2 = a_2^1 x_2$, $y^3 = a_2^2 x_2$, $y^4 = a_1 x_1 + a_2^1 x_2$, $y^5 = a_1 x_1 + a_2^2 x_2$. The main advantage of this approach is the good theoretical basis. A disadvantage is the large amount of calculations needed to maximize the utility function, which can be multi-modal if utility function is not convex. However, the main problem of this approach is reliable definition of survival probabilities. Therefore, in the next section, we implement a version of diversification defined by maximization the Sharpe ratio.

3.4.2. Trading Rule No. 6, Risk-Avoiding Users, Maximizing Sharpe Ratio in the Context of the Modern Portfolio Theory (MPT)

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

MPT models an asset's 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 a portfolio is the weighted combination of the assets' 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 1966, William Forsyth Sharpe developed what is now known as the Sharpe ratio (Sharpe, 1966). Sharpe originally called it the “reward-to-variability” ratio before it began being called the Sharpe ratio by later academics and financial operators. The definition was:

$$S = \frac{E[R - R_f]}{\sqrt{var[R]}}. \quad (3.149)$$

Sharpe’s 1994 revision (Sharpe, 1994) acknowledged that the basis of comparison should be an applicable benchmark, which changes with time. In (Sharpe, 1966) Sharpe ratio is defined as:

$$S = \frac{E[R_a - R_b]}{\sigma} = \frac{E[R_a - R_b]}{\sqrt{var[R_a - R_b]}}, \quad (3.150)$$

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 the S&P 500. $E[R_a - R_b]$ is the expected value of the excess of the asset return over the benchmark return, and σ is the standard deviation of this expected excess return.

Expected return of portfolio of assets with weights:

$$\mathrm{E}(R_p) = \sum_i w_i \, \mathrm{E}(R_i), \quad (3.151)$$

where R_p is the return on the portfolio p, R_i is the return on asset i, $w_i \geq 0$ is the weighting of component asset i (that is, the share of asset i in the portfolio), and $\sum_i w_i = 1$.

Using these symbols, the portfolio return variance can be written as:

$$\sigma_p^2 = \sum_i \sum_j w_i \, w_j \operatorname{cov}(R_i R_j). \quad (3.152)$$

Portfolio return volatility (standard deviation):

$$\sigma_p = \sqrt{\sigma_p^2}. \quad (3.153)$$

Now we estimate returns R_i of different assets i. Denote by $R_i(t)$ the return of asset i during the time interval t, for example during the day t, where $t = 1, \dots, T$. Then the sample mean is

$$\overline{R_p} = \frac{1}{T} \sum_i \sum_{t=1}^{T} w_i R_i(t), \tag{3.154}$$

and an unbiased estimator of the variance of the portfolio R_p is

$$\overline{\sigma^2} = \frac{1}{T-1} \sum_i \sum_j \sum_{t=1}^{T} w_i w_j (R_i(t) - \overline{R_i})(R_j(t) - \overline{R_j}), \tag{3.155}$$

where $R_i(t)$ is the observed profit of the i-th stock, and

$$\overline{R_i} = \frac{1}{T} \sum_{t=1}^{T} R_i(t). \tag{3.156}$$

The profits of assets $R_i(t)$ are not unique since, they depend on the particular investment procedures by different investors. Assuming that investors just keep the assets for a longer term, we can define the profit of asset i at time t as:

$$R_i(t) = r(t,i) w_i I(t), \tag{3.157}$$

where

$$r(t,i) = \beta(t,i) + \delta(t). \tag{3.158}$$

Here $I(t)$ is the funds invested at time t, $\delta(t)$ is the dividend at time t, $\beta(t,i)$ is the relative stock i price change at time t:

$$\beta(t,i) = \frac{Z(t+1,i) - Z(t,i)}{Z(t,i)}, \tag{3.159}$$

and w_i denotes the share of funds $I(t)$ invested in the asset i. In this setup, we consider the bank as an asset $i = 0$ with profitability $1 + \alpha(t)$ where $\alpha(t)$ is the bank interest at time t. We assume that the variance of this asset is zero, (the risk free asset). This information can be used to define the weights $w_i \geq 0, \sum_{0,..,p} w_i = 1$, which maximize the estimate of Sharpe ratio using standard optimization methods.

$$max_w \frac{\overline{R_p}}{\sqrt{\sigma}^2} \quad (3.160)$$

The data from time $t = 1$ until $t = T$ 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 available funds $I(t) = 1$ with corresponding adjustment of scales.

In this work, we do not consider the cyclic processes in the world finances.

3.4.3. Applying Short Term Trading Rules for the Longer Term Investment

The last four trading rules are defined by applying the short-term strategies in the longer-term investment.

We estimate the parameters by some learning set define the best portfolio according to these parameters and corresponding trading rules and buy it at the start of testing set. We sell the portfolio at the end of testing set. Note, that in the short-term environment the corresponding buying-selling operations were performed each day. We enumerate these strategies by numbers No. 7 to No. 10, accordingly.

3.5. Prediction Models

Preliminary investigation in (Mockus et al., 1997) and (Mockus and Raudys, 2010) show that the prediction of higher complexity does not necessarily provide the minimal prediction errors. This, and the limited time of simulation, suggests the preferable application of the autoregressive models, which are widely used in mathematical statistics and easily understandable intuitively. Additional advantage of the autoregressive models is the simplicity of exact solutions in the form of linear equations for AR and linear programming in AR-ABS. In contrast, for the estimation of parameters of the more advanced prediction methods, the global optimization is needed as usual.

3.5.1. AR(p) Model

Assume that the player i predicts next-day stock prices $z(t+1, i)$ using the AR(p) model (Cochrane, 2005). Professional investors are trying to obtain additional information about the fundamentals of the stock and use sophisticated statistical models. Thus the AR(p) of order p model can be regarded as a simplest simulator of a nonprofessional player which is making investments based on the data observed during past p days.

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

Assume that the stock rates changes following these simple relations

$$Z(s+1) = \sum_{k=1}^{p} a_k \, Z(s-k+1) + \varepsilon_{s+1}. \tag{3.161}$$

This formula describes the traditional autoregressive model AR(p) of order p. However, in the contest of this paper, relation (3.161) reflects opinions of stockholders that are making investment decisions based on the optimal next day predictions obtained using the past data. Later we compare the prediction models which minimize standard statistical prediction errors, such as Mean Squared Error (MSE) and Mean Absolute Error (MAE), with the models maximizing expected profit. It means that we replace the standard assumptions of the autoregressive model by the single assumption that the relation (3.161) approximately represents opinions of some stockholders.

The alternative way of fitting AR(p) parameters is the likelihood maximization which provides good mathematical results (Cochrane, 2005). However, this approach appears more difficult for stockholders intuitive understanding and the mathematical advantages are not very important regarding the AR(p) model just as a tool of the virtual stock exchange. We may consider moving average model MA(q), too, to simulate more sophisticated users which try to correct past errors, where

$$Z(s+1) = \sum_{j=1}^{q} b_j\, \varepsilon_{s-j+1} + \varepsilon_{s+1}. \qquad (3.162)$$

Minimizing the MA(q) residuals we have to minimize a polynomial function of degree t. We can see this expanding the recurrent expression (3.162). Traditional methods of parameter estimation do not consider this problem as multimodal (Cochrane, 2005). However, some more recent authors apply global optimization techniques such as particle swarm optimization (Rolf et al., 1997) and evolutionary algorithms (Voss and Feng, 2002). To represent risk-neutral users we may apply the AR-ABS model by minimizing the absolute residuals instead of the squared ones.

The PORTFOLIO model starts at time $t = 1$, so we should define the past values $Z(1-p)$. We assume that:

$$Z(s) = Z(1)(1 + \eta s), \text{ if } -p \leq s < 1. \qquad (3.163)$$

where $Z(1)$ is the initial price and $0 < \eta < 1$ is a fixed number, for example $\eta = 0{,}01$. If $1 \leq s \leq t$, then ε_s are residuals of the prediction model. Unknown parameters of AR(p) can be defined by minimization of squared residuals can be reduced to a system of linear equations and solved using efficient techniques of linear algebra.

3.5.2. AR-ABS(*p*) Model

The method of least squares is sensitive to large deviations (Arthanari and Dodge, 1993). Therefore, the replacement of squares by absolute values is beneficial, if the customers' utility function is linear. The linear utility function represents risk-neutral behavior.

The optimal prediction parameters are defined by this condition:

$$a_k^i = \arg\min_{a_k^i} \sum_{s=1}^{t} |\varepsilon_s(i)|. \qquad (3.170)$$

3.5.3. Prediction by Actual Data

In the PORTFOLIO framework, both the AR(p) and AR-ABS(p) models are meant for stock exchange simulation, assuming that stockholders predict the next-day stock prices using these models.

However, to test these models using actual data the modification separating the learning and testing procedures has been made. In the learning stage, the parameters $a_k, k = 1,\dots,p$ are estimated by expressions (3.167) using the first part of observations $1 \leq t_0 < t$. Usually t_0 is about $t/2$. During the testing stage a sequence of predictions is performed using the remaining observations $t_0 < s \leq t$ without updating the parameters. The residuals of the testing stage are used to estimate average deviations by the following expressions:

$$E_0 = \frac{1}{t - t_0 \sum_{s=t_0+1}^{t} \varepsilon_s^2(i)}, \tag{3.186}$$

or

$$E_1 = \frac{1}{t - t_0 \sum_{s=t_0+1}^{t} |\varepsilon_s(i)|}, \tag{3.187}$$

where

$$\varepsilon_s(i) = Z(s) - \sum_{k=1}^{p} a_k^i \, Z(s-k) \ - p \leq s \leq t. \tag{3.188}$$

The variance of residuals in the testing stage is estimated by this expression:

$$s^2 = \frac{\sum_{s=t_0+1}^{t} \varepsilon_s^2(i) - \frac{(\sum_{s=t_0+1}^{t} \varepsilon_s(i))^2}{t-t_0}}{t - t_0 - 1}. \tag{3.189}$$

3.6. Market Manipulation

In this section, some additional utilities are described designed to manipulate simulated financial markets.

3.6.1. Forcing Sells and Buys

To force-sell in order to depress prices, we need to set the low selling level. To force-buy, one sets high buying levels. To normalize prices, we should restore normal buying-selling levels or to set the new buying levels. These operations can be conveniently performed during the 'Stop' mode.

In the 'force sell' window, we set the value of the multiplier ν_s. This multiplier defines how much the new selling level $z_{fs}(t,i,l)$ is depressed in the interval between the minimal and normal selling levels. The buying levels are disabled during force-sell operation.

In the window 'force-sell', we set the value of the multiplier ν_s defining the force-sell mode. In the window 'force-buy', we set the value of the multiplier ν_b defining the force-buy mode.

In the 'force-sell' mode, using (3.2), (3.42) and (3.44), we write the modified sell and buy levels in this form:

$$\begin{aligned} z_s^s(t,i,j,l) &= \nu_s z_s(t,i,j,l), \\ z_b^s(t,i,j,l) &= \nu_s z_b(t,i,j,l), l = 1,2,3, \end{aligned} \tag{3.198}$$

where $\delta + 3\tau_0 \leq \nu_s \leq 1$.

In the 'force-buy' mode, the sell and buy levels are as follows:

$$\begin{aligned} z_s^b(t,i,j,l) &= \nu_b z_s(t,i,j,l), \\ z_b^b(t,i,j,l) &= \nu_b z_b(t,i,j,l), l = 1,2,3, \end{aligned} \tag{3.199}$$

where $\nu_b \geq \delta + 3\tau_0$, the default is $\nu_s = 1{,}0$ and $\nu_b = 1{,}0$.

It follows from (3.44) and (3.196) that in the 'force-sell' mode, the buying-selling profitability levels can be defined this way

$$\begin{aligned} p_s^s(t,i,j,l) &= \frac{z(t+1,i,j)}{z_s^s(t,i,j,l)} - \big(1 - \delta(t) + \alpha(t) + h(t)\big), \\ p_b^s(t,i,j,l) &= \frac{z(t+1,i,j)}{z_b^s(t,i,j,l)} - \big(1 - \delta(t) + \alpha(t) + h(t)\big). \end{aligned} \tag{3.200}$$

Here $0 < z_{fs}(t,i,j,l) \leq z_{fb}(t,i,j,l) < \infty$.

In the 'force-buy' mode, the buying-selling profitability levels can be defined this way

$$p_s^b(t,i,j,l) = \frac{z(t+1,i,j)}{z_s^b(t,i,j,l)} - \big(1 - \delta(t) + \alpha(t) + h(t)\big),$$
$$p_b^b(t,i,j,l) = \frac{z(t+1,i,j)}{z_b^b(t,i,j,l)} - \big(1 - \delta(t) + \alpha(t) + h(t)\big). \tag{3.201}$$

3.7. Conclusions of Chapter 3

The PORTFOLIO model provides the possibilities to simulate the stock exchange processes in the multi-stock and multi-user environment in both the real and virtual markets where the stock prices are generated by the interaction of different investors using different trading rules and different investment models. Apparently, this is the first model including all these features together.

4

Experimental Research

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

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

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

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

Period III. 352 working days from 2013-07-19, this experiment shows the present times.

The historical data of the following eight stocks of companies was 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 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.

Thus, in all these simulations, a sub set of 80 trading strategies (selected from the set of 190 strategies) were investigated. We define the trading strategy as a pair of trading rule and prediction model.

The objectives of the experiments are to investigate in both the historical and virtual environments:

1. The relation of profits on prediction errors using different investment strategies at different economic conditions.
2. The relation of profits on different trading rules and prediction models.
3. The relation of optimal portfolios on different investment strategies.

The complete experimental results are presented in tables. In addition, some selected results are illustrated by column-charts. To illustrate the most important unexpected result, the correlations of average profits and prediction errors are calculated and presented in the form of column-charts with confidence intervals. The specific properties of different portfolios are illustrated as pie-charts. The results are self-explanatory, so only the minimal comments are written.

4.1. Real Stock Experiment – Period I

In this section, the experiment with historical data of the first period is discussed.

In Table 1, average profits of ten trading rules and eight prediction modes are presented. In this recovery period, the maximal profit (20258.18) was achieved by trading rule No. 1 (TR1) and AR(6). The greatest loss (-3013.07) occurred using TR3 and AR-ABS(6). The corresponding portfolios are in Table 4.

Table 1 *Average profits of eight prediction modes and ten trading rules in real stock market, Period I*

Trading Rule	AR-ABS(1)	AR-ABS(3)	AR-ABS(6)	AR-ABS(9)	AR(1)	AR(3)	AR(6)	AR(9)
TR1	1059.36	2565.77	-518.07	7151.94	3864.07	3340.51	**20258.18**	7669.38
TR2	6181.79	-524.03	-2334.40	4908.25	5291.30	589.08	1962.95	525.94
TR3	1287.35	-431.64	**-3013.07**	-2459.46	42.58	-1169.67	8182.93	2665.21
TR4	3806.22	3777.34	5629.49	7151.88	5579.59	4114.89	2456.41	4637.71
TR5	98.32	255.59	76.36	80.44	76.61	84.03	55.19	126.96
TR6	104.74	172.13	106.81	147.21	78.54	155.51	76.37	169.17
TR7	215.75	246.59	226.74	-151.71	275.74	182.30	456.44	-118.56
TR8	157.95	189.20	125.61	187.59	175.38	99.35	502.59	158.26
TR9	152.44	276.33	145.04	-19.42	292.92	55.86	178.46	-239.02
TR10	212.88	217.46	224.92	127.27	153.77	154.98	278.39	73.39

In Tables 2 and 3, the Mean Absolute Error (MAE) and the Standard Error (SE) of eight prediction modes are presented. Here one can see that the greatest prediction error was using the AR(9) model. Other prediction models show smaller prediction errors. Some insignificant differences between the prediction errors obtained by different trading rules can be explained by a shift in the starting point due to technical reasons. In contrast, the differences of profits obtained by different trading rules are significant in this and the following time periods, see Figures 5, 9 and 14.

Table 2 *MAE in real stock market, average of eight stocks, Period I*

Trading Rule	MAE							
	AR-ABS(1)	AR-ABS(3)	AR-ABS(6)	AR-ABS(9)	AR(1)	AR(3)	AR(6)	AR(9)
TR1	0.01844	0.01845	0.01861	0.01864	0.01829	0.02398	0.029054	0.0327
TR2	0.01844	0.01845	0.01861	0.01864	0.01829	0.02398	0.029054	0.0327
TR3	0.01844	0.01845	0.01861	0.01864	0.01829	0.02398	0.029054	0.0327
TR4	0.0184	0.01844	0.01857	0.0186	0.01825	0.02043	0.025877	0.0314
TR5	0.01959	0.0196	0.01966	0.01971	0.01954	0.04167	0.030465	0.06659
TR6	0.01959	0.0196	0.01966	0.01971	0.01954	0.04167	0.030465	0.06659
TR7	0.01949	0.01951	0.01959	0.01961	0.01945	0.03942	0.062136	0.06901
TR8	0.0195	0.01953	0.01965	0.01966	0.01944	0.04742	0.062004	0.0815
TR9	0.01951	0.01946	0.01953	0.01953	0.0194	0.02249	0.029292	0.03585

TR10	0.01944	0.0195	0.01953	0.01964	0.01938	0.02905	0.027745	0.14727

Table 3 *SE in real stock market, average of eight stocks, Period I*

Trading Rule	SE							
	AR-ABS(1)	**AR-ABS(3)**	**AR-ABS(6)**	**AR-ABS(9)**	**AR(1)**	**AR(3)**	**AR(6)**	**AR(9)**
TR1	0.00232	0.00232	0.00233	0.00233	0.00232	0.00639	0.007564	0.00618
TR2	0.00232	0.00232	0.00233	0.00233	0.00232	0.00639	0.007564	0.00618
TR3	0.00232	0.00232	0.00233	0.00233	0.00232	0.00639	0.007564	0.00618
TR4	0.00232	0.00232	0.00233	0.00232	0.00231	0.00279	0.004348	0.00591
TR5	0.00265	0.00265	0.00265	0.00265	0.00264	0.02231	0.007393	0.03649
TR6	0.00265	0.00265	0.00265	0.00265	0.00264	0.02231	0.007393	0.03649
TR7	0.00259	0.00259	0.0026	0.0026	0.00259	0.02031	0.037425	0.03139
TR8	0.0026	0.0026	0.0026	0.0026	0.0026	0.0283	0.03706	0.03454
TR9	0.00255	0.00255	0.00255	0.00255	0.00255	0.00385	0.006043	0.00847
TR10	0.00252	0.00253	0.00253	0.00253	0.00252	0.01022	0.005292	0.09705

Table 4 shows best average portfolios, using ten said trading rules and eight prediction modes. The most profitable portfolio TR1 and AR(6) contains mainly BAC stocks. The explanation is the rapid recovery of the BAC stock prices in the post-crisis period. Another reason is no diversification, since TR1 is more risky as compared with other trading rules used in this research.

Table 4 *Average portfolios of ten trading rules in real stock market, Period I*

Stock name	TR1	TR2	TR3	TR4	TR5
	AR(6)	**AR-ABS(1)**	**AR(6)**	**AR-ABS(9)**	**AR-ABS(3)**
MSFT	37.68	25.393	50.23	193.31	5.51
AAPL	0.66	86.864	24.41	49.24	0.27
GOOG	0.36	0.000	2.00	3.28	0.04
NOK	7.24	0.054	121.13	25.42	0.23
TM	1.28	0.000	9.38	2.81	0.01
BAC	1937.03	6.044	496.08	65.03	7.68
BA	1.14	0.007	24.96	12.33	0.03
ORCL	43.43	0.086	43.41	10.77	4.10
AVG (strategy)	**253.60**	14.8062	96.45	45.2759125	**2.23**
Stock name	**TR6**	**TR7**	**TR8**	**TR9**	**TR10**
	AR-ABS(3)	**AR(6)**	**AR(6)**	**AR(1)**	**AR(6)**
MSFT	3.5754	5.8199	1.88	0	2.7757

AAPL	0.3973	0.6925	0.64	3.1373	1.4073
GOOG	0.2323	0.0135	1.15	0.2484	0.3728
NOK	0.4553	7.3817	1.96	0	1.7831
TM	0.1024	0.5652	0.67	0.6212	1.0178
BAC	3.0609	14.6757	4.23	0.2486	2.4461
BA	0.1599	0.2494	0.43	0.1738	0.8598
ORCL	2.6403	9.2168	1.08	0	3.2213
AVG (strategy)	**1.33**	**4.83**	**1.51**	**0.55**	**1.74**

Figure 3 presents average forecast errors of 8 traders using TR1 and different forecast methods. TR1 was selected since it provided the greatest profit. The chart show that in this conditions, the greatest error occurred using the most complicated prediction model AR(9).

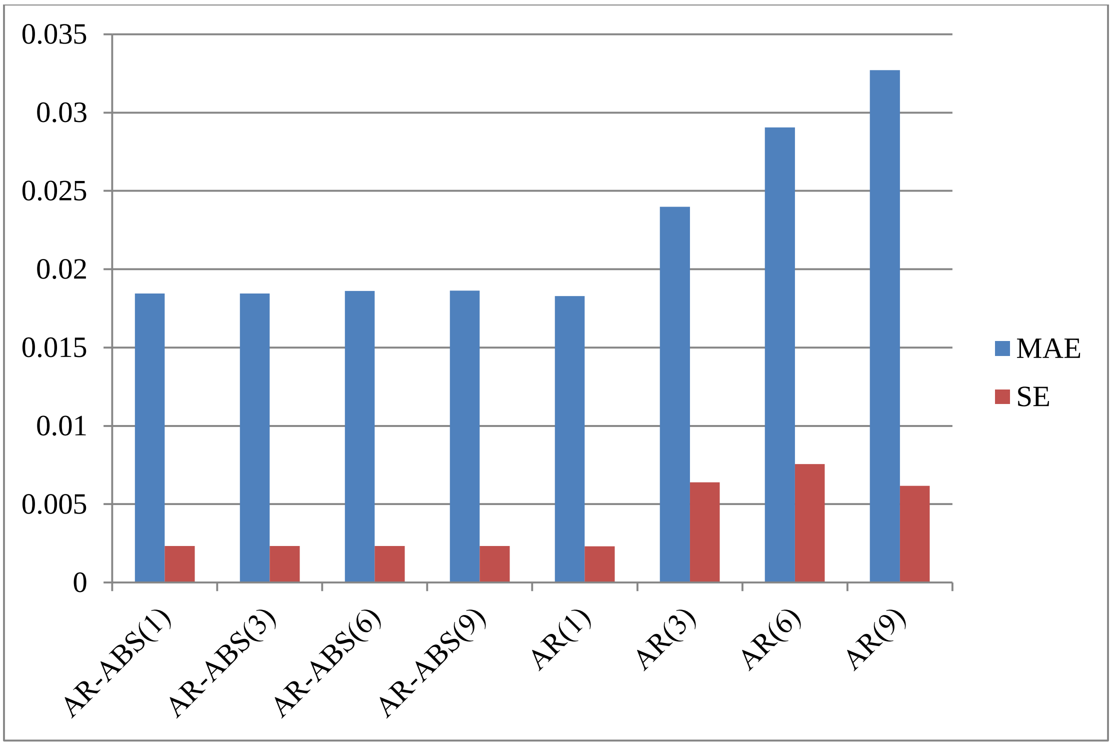

Fig. 3 *MAE and SE in real stock market, average of eight stocks, Period I, using TR1*

Comparison of Figures 3 and 5 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 which prediction error is close to the maximal one. This paradoxical situation is confirmed by the positive profit-prediction error correlations in Figure 34. Figure 4 illustrates irregular growth of stock prices in the post-crisis economical conditions.

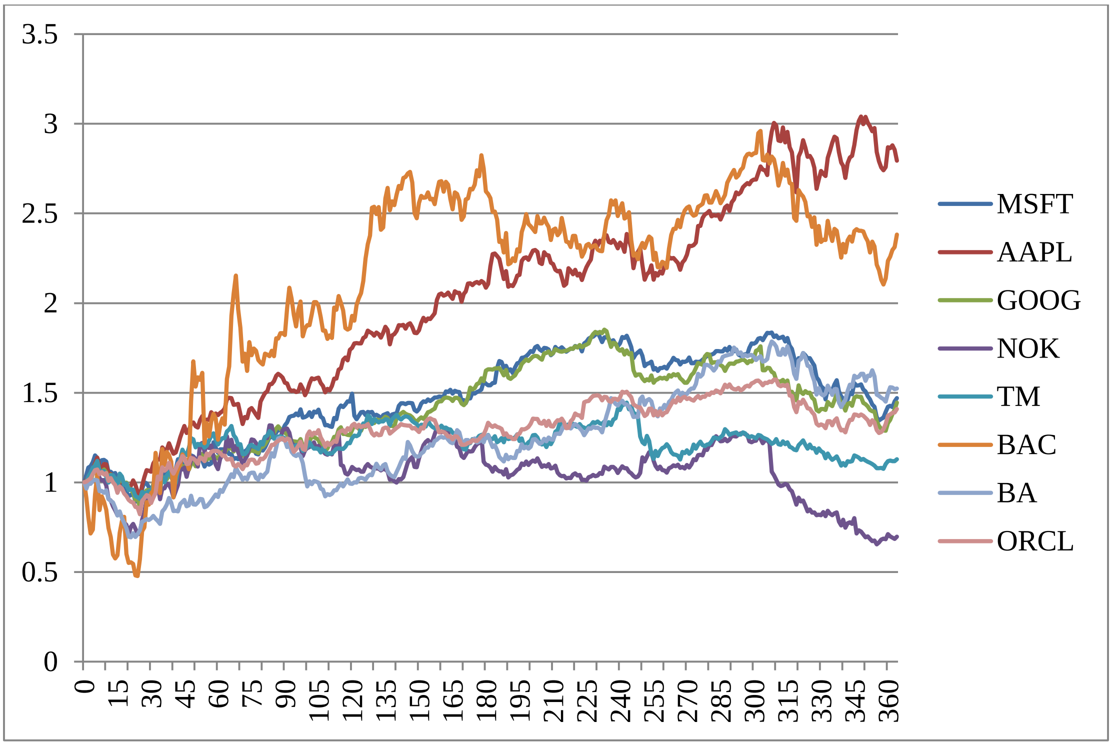

Fig. 4 *Normalized daily prices of eight stocks in the post-crisis Period I*

Interesting observation is the positive correlations were in the both irregular growth periods: one is this post-crisis period, another is generated by virtual stock exchange, see Figure 28 where the stock price graph is similar to the post-crisis period. The corresponding profit-prediction correlations are presented in Figures 34 and 37.

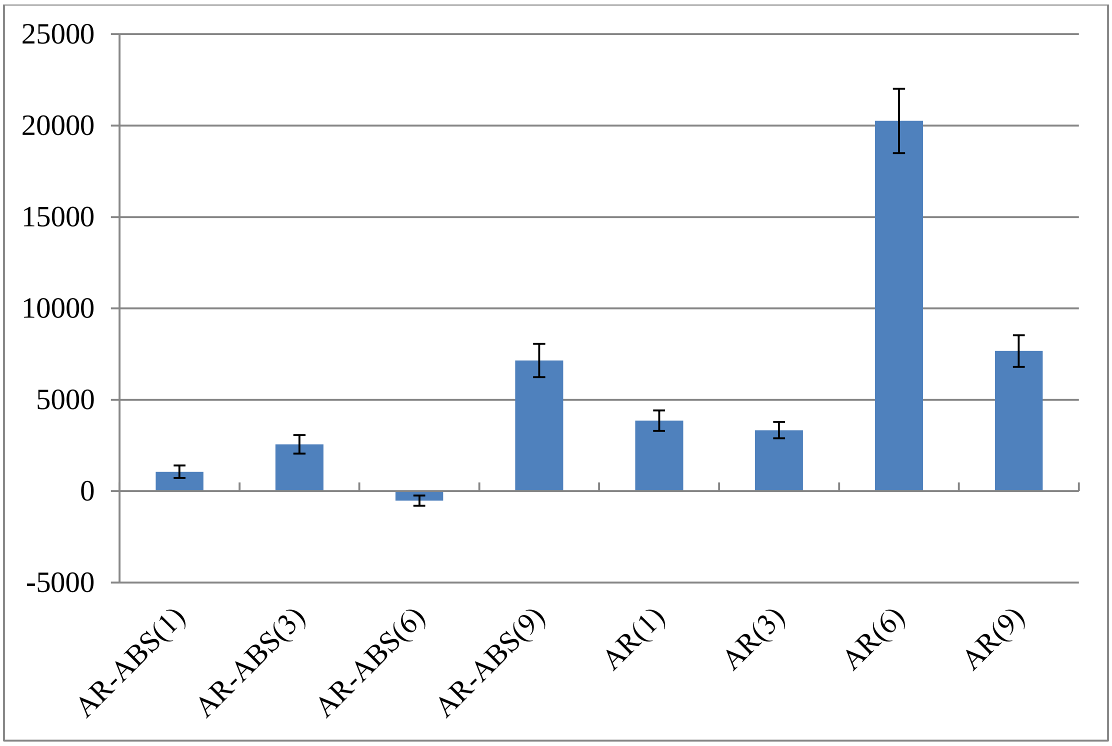

Fig. 5 *Average profits of eight prediction modes in real stock market, Period I, using TR1*

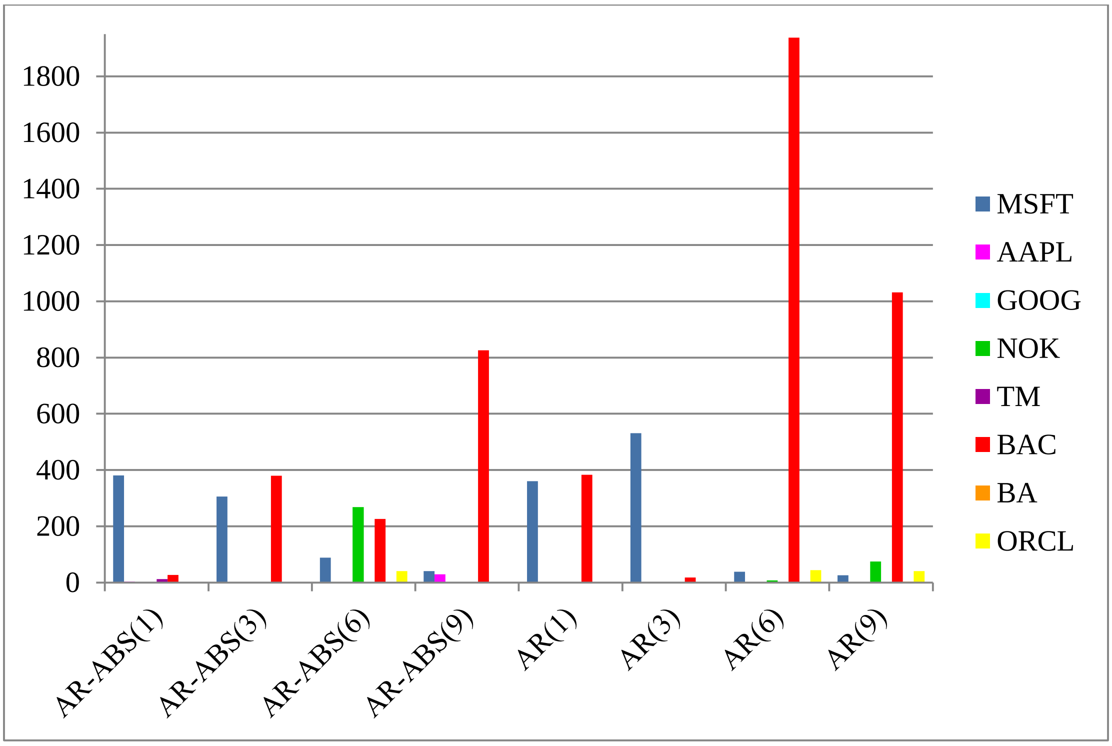

Fig. 6 *Average portfolios in real stock market, Period I, using TR1 and different prediction modes*

Figure 6 shows average portfolios using different forecast methods and TR1.

One can see, that stocks in portfolio distributed unevenly. Using trading TR1 and different forecast methods traders mostly preferred BAC and MSFT stocks. Also NOK, ORCL and AAPL stocks were traded. All other stocks were traded less or even not traded at all.

In the first period, the greatest profit was obtained using TR1 and the forecast model AR(6). The corresponding average portfolio contains mostly the BAC stock.

Figure 7 shows the graph of portfolios of investors using TR4 and prediction model AR(1).

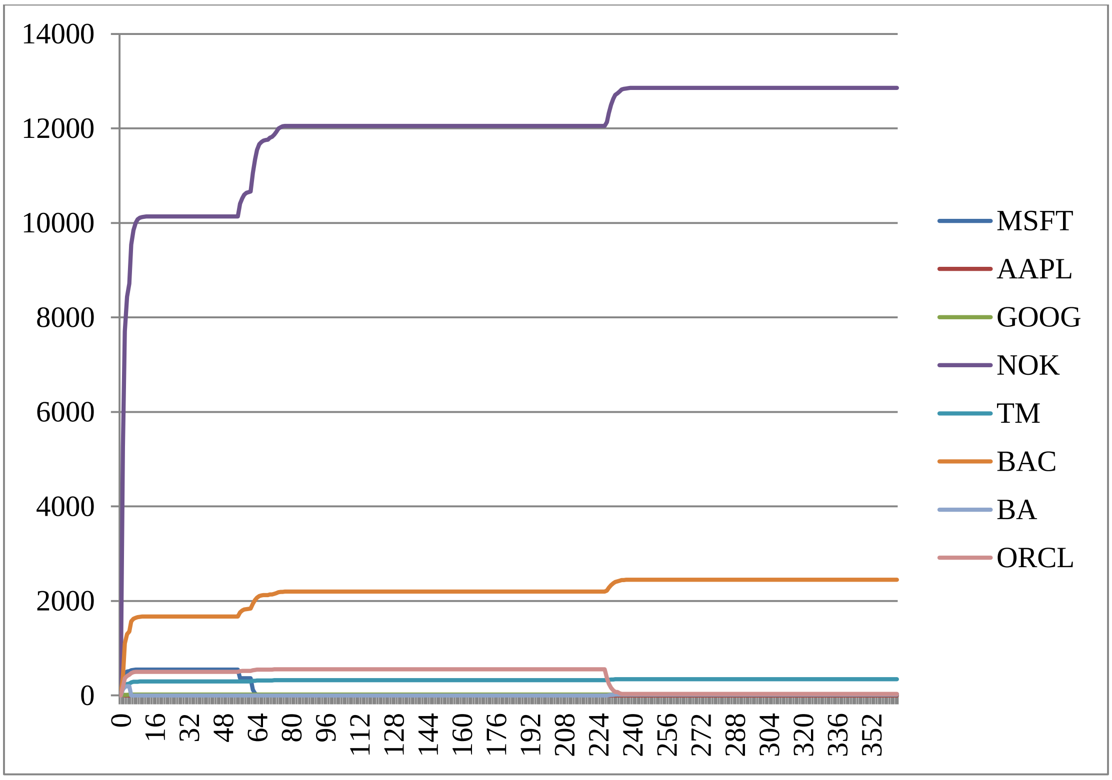

Fig. 7 *Portfolio graph in real stock market, Period I, using TR4 and AR(1)*

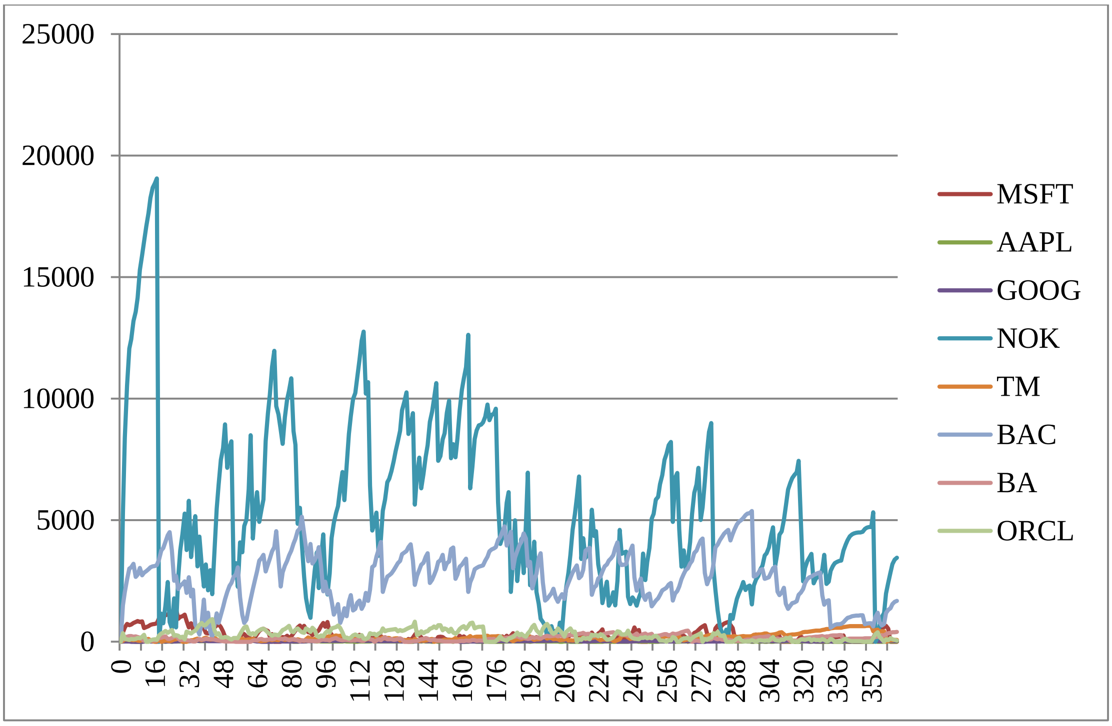

Fig. 8 *Portfolio graph in real stock market, Period I, using TR4 and AR(9)*

Comparing Figures 7 and 8 one can see that using prediction model AR(9) the trading is irregular and different from the corresponding 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, see Table 1.

To illustrate the sensitivity of the results to trading rules, TR6 is considered for comparison. Figure 9 shows that MAE and SE of eight prediction modes, using TR6 follows similar pattern as in TR1, as expected.

In contrast, the charts of profits in Figures 5 and 10 are different. The corresponding confidence intervals show that the differences are significant statistically. This illustrates that, in this experiment, the profits depended mainly on the trading strategies. The correlation of profit to prediction accuracy is negative as often as positive, see Figure 34.

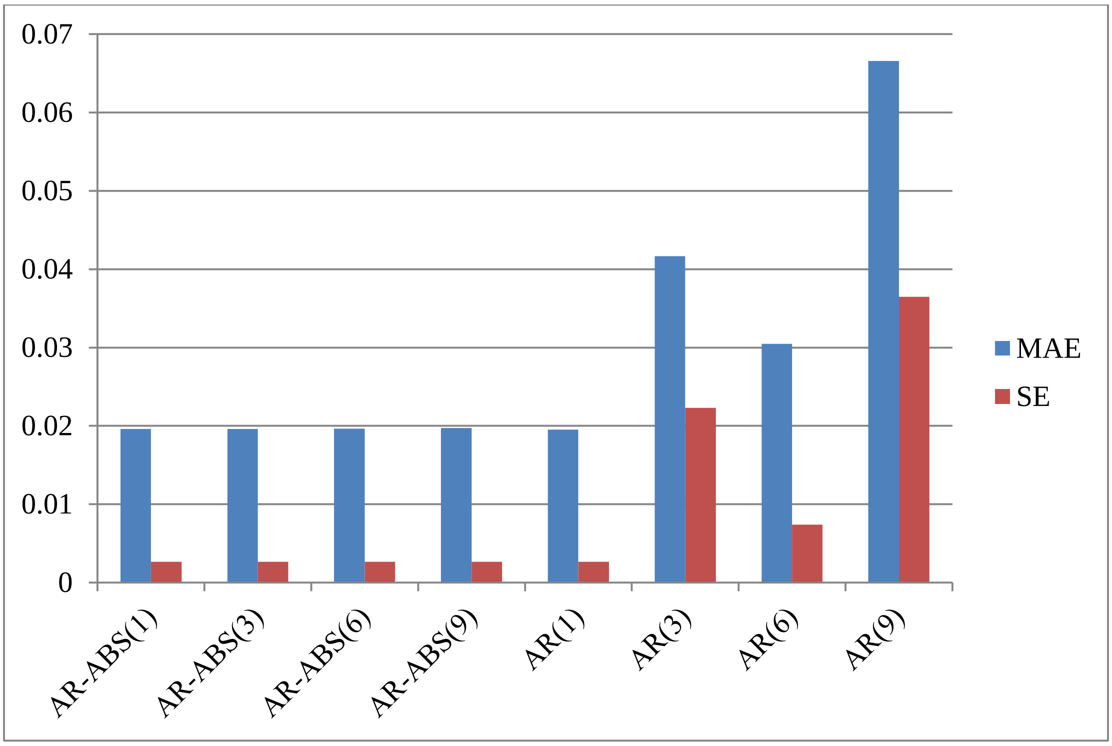

Fig. 9 *MAE and SE in real stock market, average of eight stocks, Period I, using TR6*

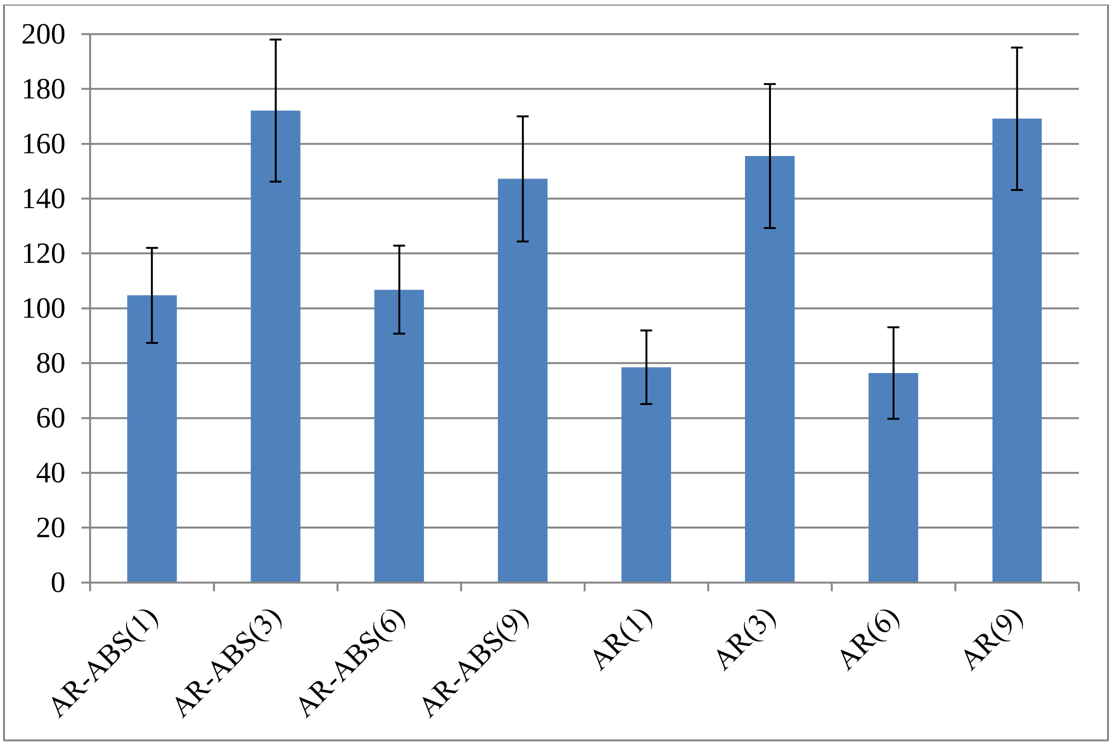

Fig. 10 *Average profits of eight prediction modes in real stock market, Period I, using TR6*

Figure 11 presents average portfolios of eight prediction modes, using TR6. These portfolios differ from those that were obtained using TR1, they are more diversified but provide lesser profits. Here BAC and MSFT stocks are the most popular, but, unlike the previous case, ORCL stocks are also presented.

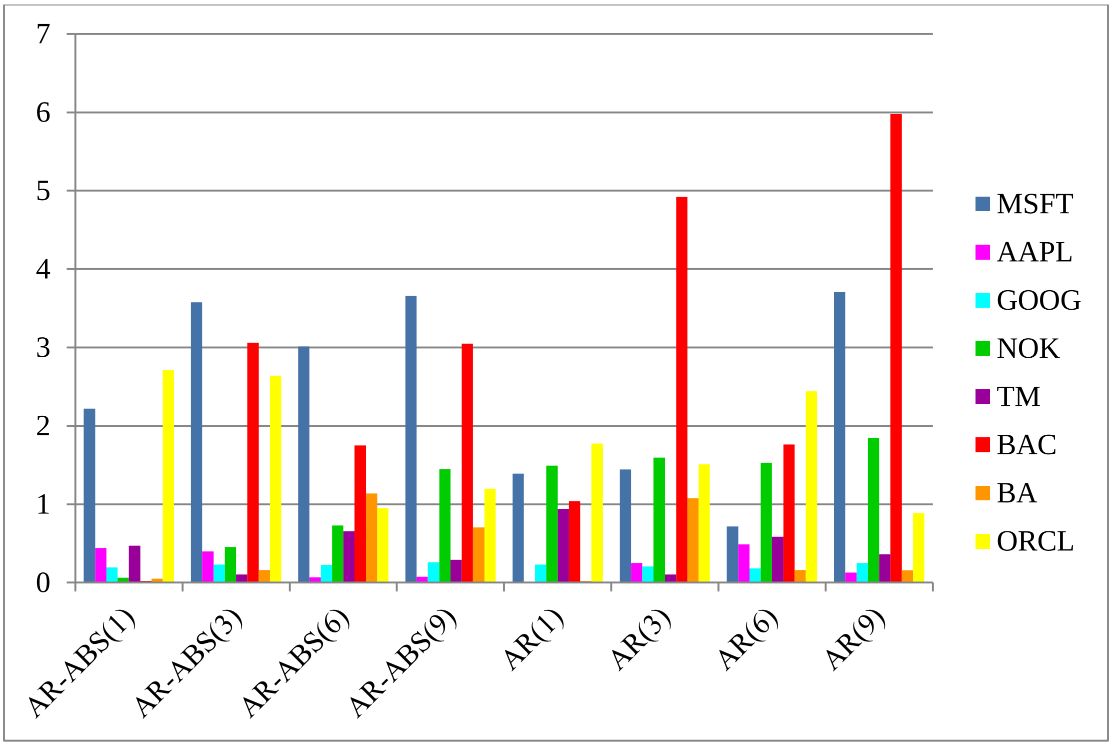

Fig. 11 *Average portfolios in real stock market, Period I, using TR6 and different prediction modes*

In the Period I, using the risk-avoiding TR6, the greatest profit was obtained by the more diversified portfolio of the forecast model AR-ABS(3). Mostly MSFT, BAC, ORCL were included.

However, the profit of this portfolio was just 172.13 as compared with the profit 20258.18 of the best portfolio defined by the more risky TR1. This illustrates the cost of risk avoiding.

Figure 12 illustrates the longer time trading process where investors are buying stocks at the end of learning period and selling them at the end of testing period.

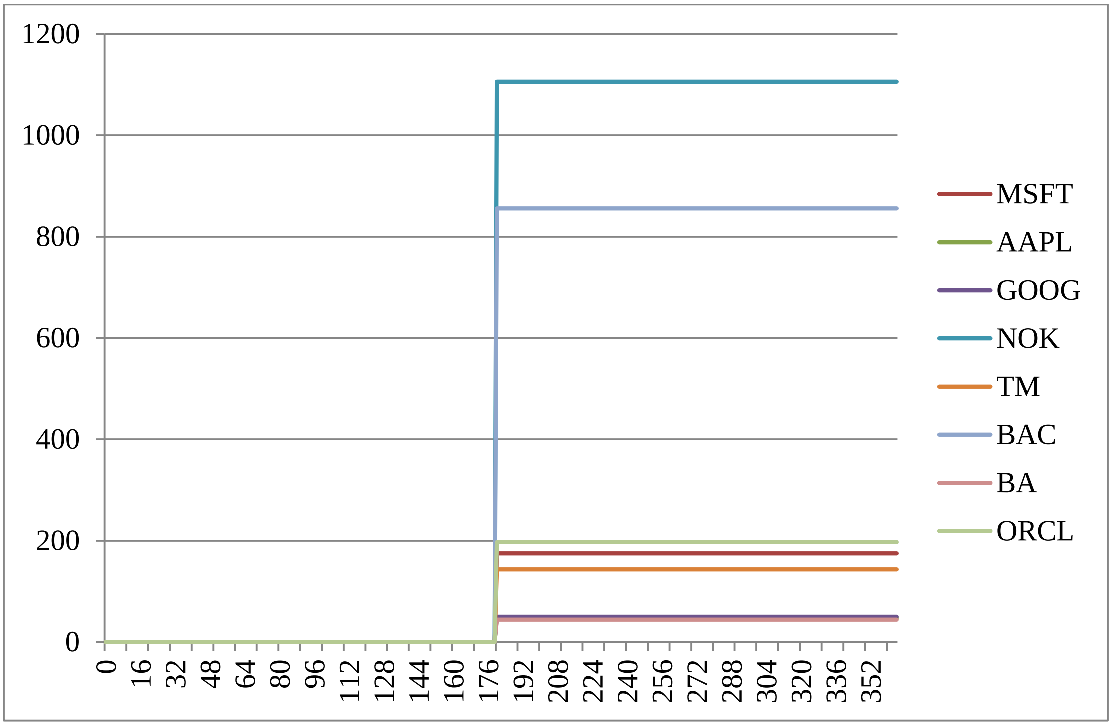

Fig. 12 *Portfolio graph in real stock market, Period I, using TR6*

4.2. Real Stock Experiment – Period II

In this section, the results of the second, more stable period are presented. Table 5 shows profits of eight prediction modes and ten trading rules. The greatest profit was obtained by TR1 and AR-ABS(1). The greatest losses show TR7 and AR(3).

Table 5 *Average profits of eight prediction modes and ten trading rules in real stock market, Period II*

Trading Rule	AR-ABS(1)	AR-ABS(3)	AR-ABS(6)	AR-ABS(9)	AR(1)	AR(3)	AR(6)	AR(9)
TR1	**22916.86**	22821.92	22824.83	22855.84	22828.96	-7002.53	-36921.4	-3982.17
TR2	5916.725	16270.38	1596.154	3961.94	2972.32	-16364.4	-15633.4	-7023.07
TR3	8501.064	13992.02	-15252.1	-16964.2	212.3845	-18251.7	-36145.5	-32364.4
TR4	-5416.18	4121.961	9658.151	6570.48	-5375.21	-92.954	-357.497	3003.661
TR5	1392.371	1381.205	1084.238	1553.303	1164.452	1416.071	1241.34	1706.036
TR6	5644.673	5683.615	5648.979	5796.353	6019.915	5707.711	5708.92	5687.203
TR7	-7220.31	-7309.89	3041.269	3016.672	-7404.73	**-50152.8**	-17276	6200.664

TR8	6636.983	9970.032	5925.754	-6786.48	-1924.34	-15383.4	-1698.71	-5510.49
TR9	-17219.2	-8951.79	-8345.76	-19836.1	-1070.85	-14672.4	-32388.1	-11834.6
TR10	8000.388	7438.16	-151.467	-6042.18	-3007.51	-6924.21	-4619.83	-3267.78

Tables 6 and 7 show MAE and SE obtained by eight different prediction modes and ten trading rules. Comparing Tables 6 and 7 with the Period I MAE and SE Tables 2 and 3 we see similar pattern. Some numerical differences are not as great as expected due to different economic conditions.

TR6 and TR7 are exceptions; both the patterns and values of prediction errors are different not only from Period I, but also from other trading rules of Period II. A possible explanation is numerical instability of AR models, which are sensitive to small data changes. The AR-ABS models are more stable, so no unexpected differences were observed using these models.

This instability can be explained by greater sensitivity of AR(p) models to seemingly insignificant differences in time series as compared with AR-ABS(p) models, especially at larger p. The reason is that at some data, the system of linear equations minimizing the squared deviation becomes ill-defined computationally (determinant close to zero). Minimizing the absolute deviations in the AR-ABS(p) models, one uses Linear Programming which is less sensitive. Note, that errors of AR-ABS models are similar to AR(p) models at small p (up to $p = 3$) because the corresponding systems of just one, two or three linear equation are well defined, as usual.

Table 6 *MAE in real stock market, average of eight stocks, Period II*

Trading Rule	**MAE**							
	AR-ABS(1)	**AR-ABS(3)**	**AR-ABS(6)**	**AR-ABS(9)**	**AR(1)**	**AR(3)**	**AR(6)**	**AR(9)**
TR1	0.016299	0.016317	0.016275	0.016291	0.016240	0.017668	0.023934	0.036647
TR2	0.016322	0.016334	0.016309	0.016304	0.016261	0.021973	0.027580	0.071546
TR3	0.016360	0.016331	0.016340	0.016365	0.016294	0.019264	0.060082	0.032156
TR4	0.016239	0.016271	0.016277	0.016284	0.016209	0.018895	0.036418	0.029910
TR5	0.016239	0.016271	0.016277	0.016284	0.016209	0.018895	0.036418	0.029910
TR6	0.016284	0.016280	0.016298	0.016293	0.016233	0.020702	0.216166	0.071849

TR7	0.016284	0.016280	0.016298	0.016293	0.016233	0.020702	0.216166	0.071849
TR8	0.016310	0.016277	0.016315	0.016309	0.016295	0.019525	0.036981	0.035999
TR9	0.016236	0.016280	0.016270	0.016322	0.016244	0.019184	0.022766	0.024683
TR10	0.016237	0.016267	0.016284	0.016305	0.016235	0.018114	5.172787	0.038685

In Table 7, SE of eight different prediction modes and ten trading rules are shown.

Table 7 *SE in real stock market, average of eight stocks, Period II*

Trading Rule	**SE**							
	AR-ABS(1)	**AR-ABS(3)**	**AR-ABS(6)**	**AR-ABS(9)**	**AR(1)**	**AR(3)**	**AR(6)**	**AR(9)**
TR1	0.002825	0.002826	0.002825	0.002825	0.002825	0.003055	0.00516	0.012445
TR2	0.002853	0.002853	0.002853	0.002852	0.002852	0.006174	0.00745	0.04577
TR3	0.002856	0.002856	0.002856	0.002856	0.002856	0.003748	0.03756	0.007339
TR4	0.002819	0.00282	0.002819	0.002819	0.002818	0.003514	0.01416	0.006983
TR5	0.002819	0.00282	0.002819	0.002819	0.002818	0.003514	0.01416	0.006983
TR6	0.002845	0.002845	0.002845	0.002845	0.002845	0.005113	0.193	0.037534
TR7	0.002845	0.002845	0.002845	0.002845	0.002845	0.005113	0.193	0.037534
TR8	0.002844	0.002843	0.002843	0.002843	0.002843	0.00377	0.01429	0.009387
TR9	0.002874	0.002874	0.002875	0.002875	0.002875	0.00383	0.00501	0.004512
TR10	0.002891	0.002892	0.002893	0.002893	0.002892	0.003312	5.15504	0.013779

Table 8 shows the best average portfolios of eight different prediction modes and ten trading rules. Most profitable portfolio includes just one company, namely BAC. Other profitable portfolios also preferred BAC stocks, but included some ORCL, AAPL, GOOG, NOK and MSFT stocks, too.

Table 8 *Average portfolios of ten trading rules in real stock market, Period II*

Stock name	**TR1**	**TR2**	**TR3**	**TR4**	**TR5**
	AR-ABS1	**AR-ABS3**	**AR-ABS3**	**AR-ABS6**	**AR9**
MSFT	0	0	8.8575	61.8137	85.3373
AAPL	0	69.7452	36.7726	28.9945	13.7268
GOOG	0	74.5425	10.6767	47.6904	2.3474
NOK	0	0	426.9096	170.2027	126.7989
TM	0	19.0767	32.7041	19.9205	34.2049
BAC	12692.2658	1374.1644	5800.0493	1241.1534	449.1986
BA	0	10.8904	96.4329	53.3699	26.8521

ORCL	0	59.4959	112.8027	1063.0274	135.9518
AVG (strategy)	**1586.53**	**200.99**	**815.65**	**335.77**	**109.30**
Stock name	**TR6**	**TR7**	**TR8**	**TR9**	**TR10**
	AR(1)	**AR(9)**	**AR-ABS(3)**	**AR(1)**	**AR-ABS(1)**
MSFT	56.7321	98.4055	0	0	30.2767
AAPL	11.9811	4.2	67.5342	78.9151	12.2521
GOOG	13.6849	0	0	0	6.3452
NOK	636.5047	3.0658	150.3534	248.5753	16.3452
TM	22.177	0.0356	0	3.274	3.137
BAC	220.6775	5330.7014	778.6767	180.2219	3457.6192
BA	20.0293	6.1507	14.6192	0	14.9123
ORCL	63.3945	18.6384	81.5288	0	245.9644
AVG (strategy)	**130.65**	**682.65**	**136.59**	**63.87**	**473.36**

Figure 13 presents average forecast errors of 8 traders by different forecast methods, using TR1. The pattern is similar to the Period I, see Figure 9. The differences in the numerical values are less than expected, since economic conditions differ.

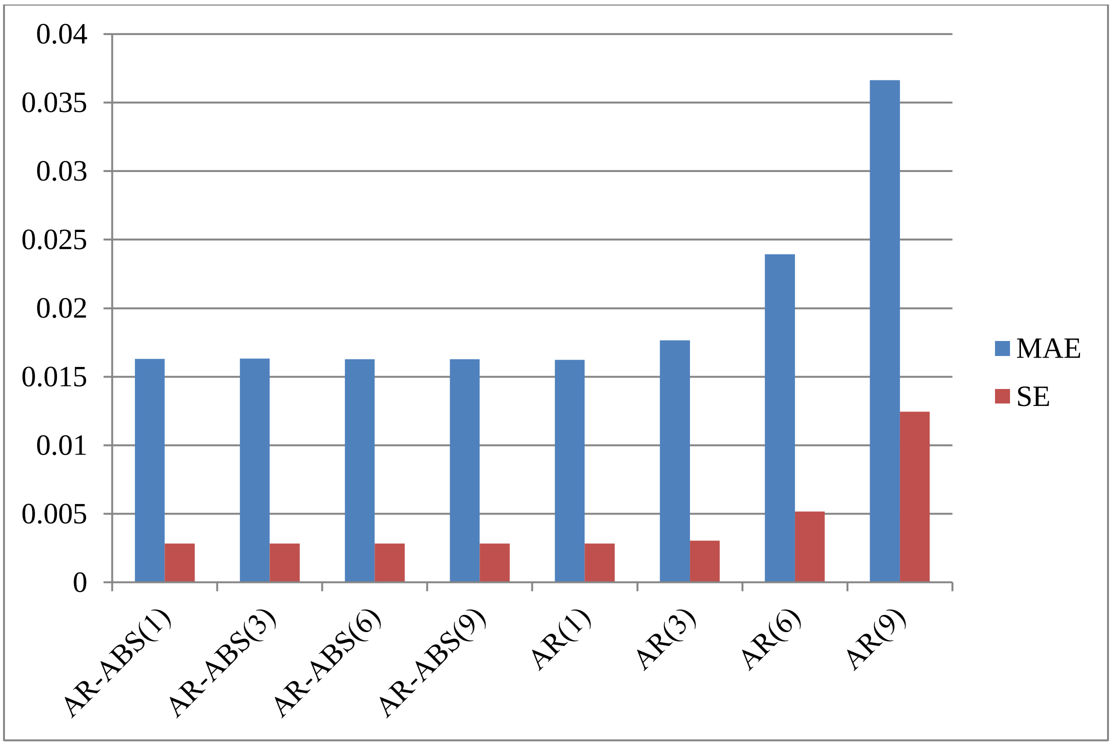

Fig. 13 *MAE and SE in real stock market, average of eight stocks, Period II, using TR1*

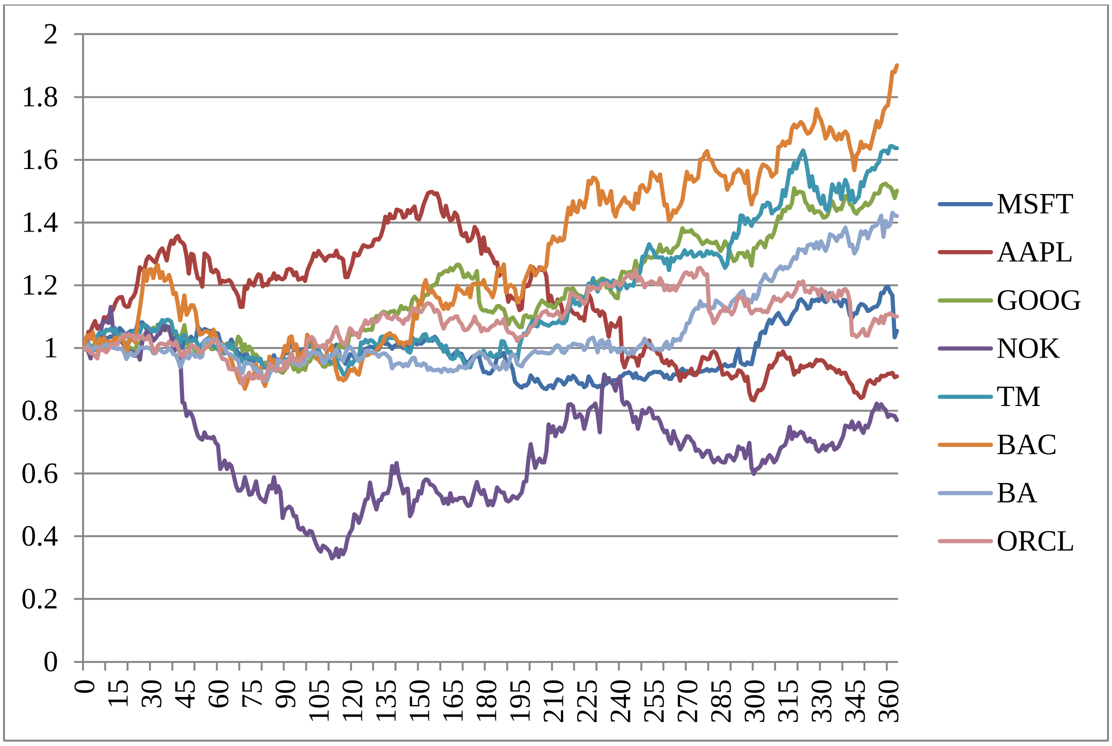

Fig. 14 *Normalized daily prices of eight stocks in Period II*

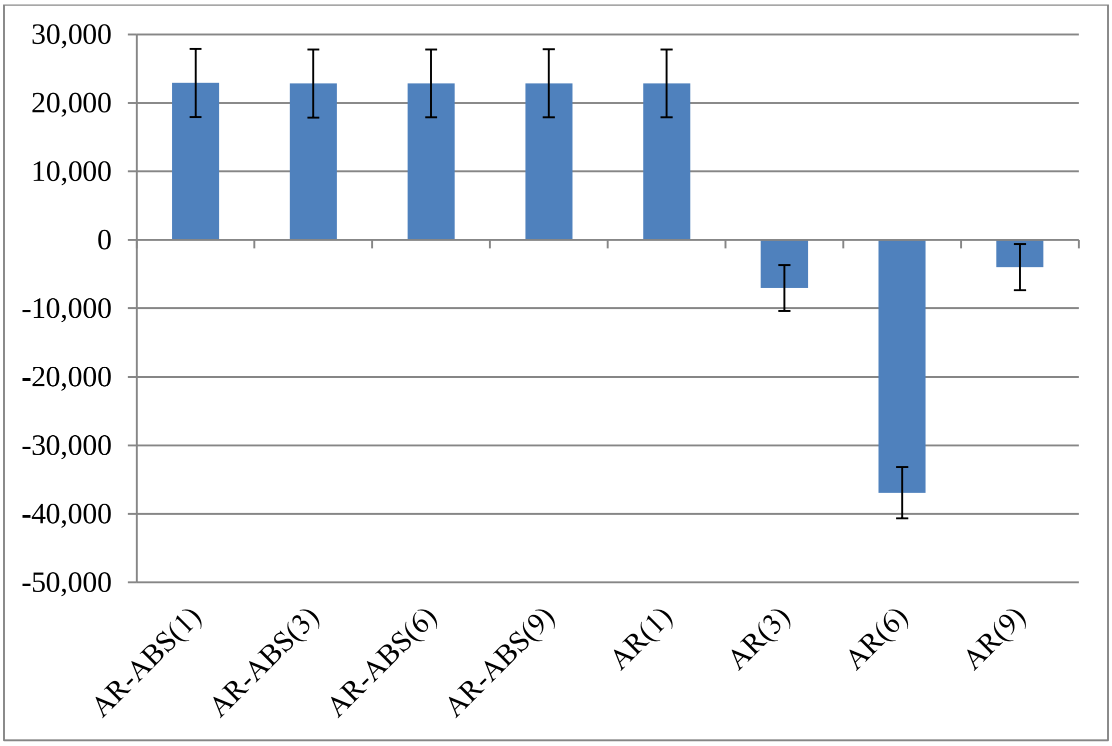

Fig. 15 *Average profits of eight prediction modes in real stock market, Period II, using TR1*

Figure 14 shows the normalized daily stock prices in the Period II when the market conditions was more stable.

Figure 15 shows average profits of eight prediction modes, using TR1.

The pattern of profits is different from the first period. Here, the most profitable are five strategies: 1) TR1 and AR-ABS(1); 2) TR1 and AR-ABS(3); 3) TR1 and AR-ABS(6); 4) TR1 and AR-ABS(9); 5) TR1 and AR(1). Their profits are almost identical. The remaining three strategies show losses and the biggest loss happened by the same TR1 and AR(6), which provided the greatest profit in Period I, see Figure 5.

Figure 16 shows average portfolios by eight prediction modes, using TR1.

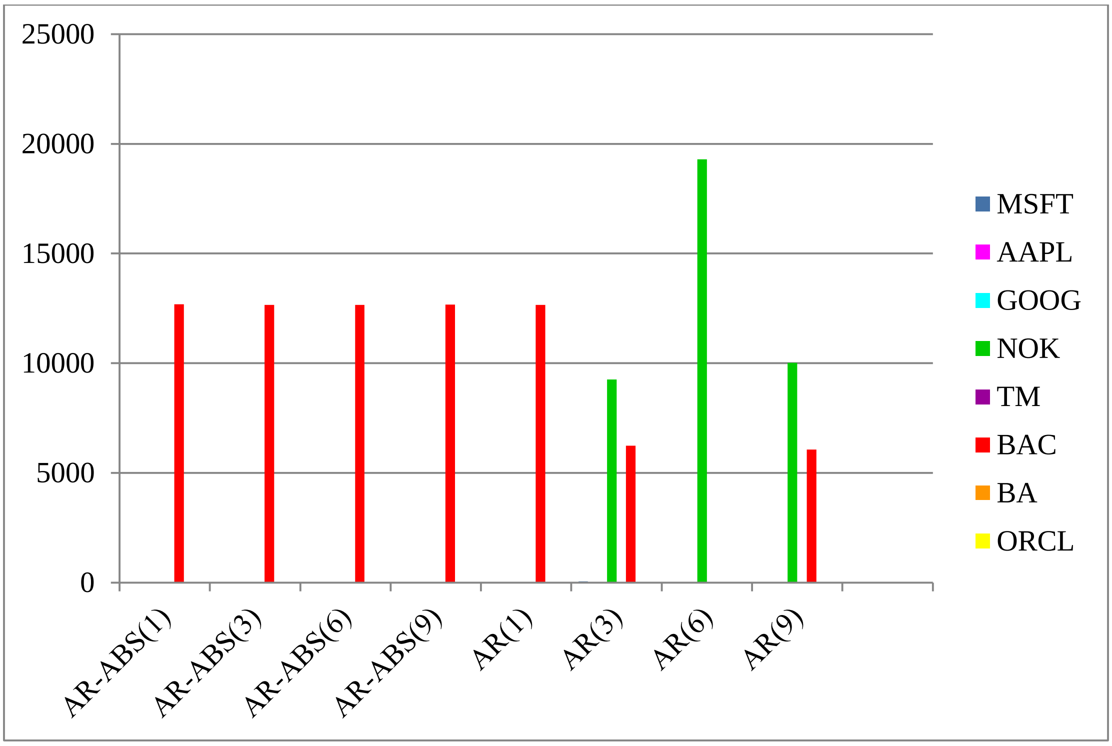

Fig. 16 *Average portfolios in real stock market, Period II, using TR1 and different prediction modes*

Only two stocks BAC and NOK are included, the others were ignored. All losing strategies included NOK stocks. The strategy with biggest losses included only NOC stocks.

In Period II, the greatest profit was obtained using TR1 and the forecast model AR-ABS(1). The best average portfolio contains only BAC stocks. Figure 17 shows average forecast errors of eight prediction modes, using TR6.

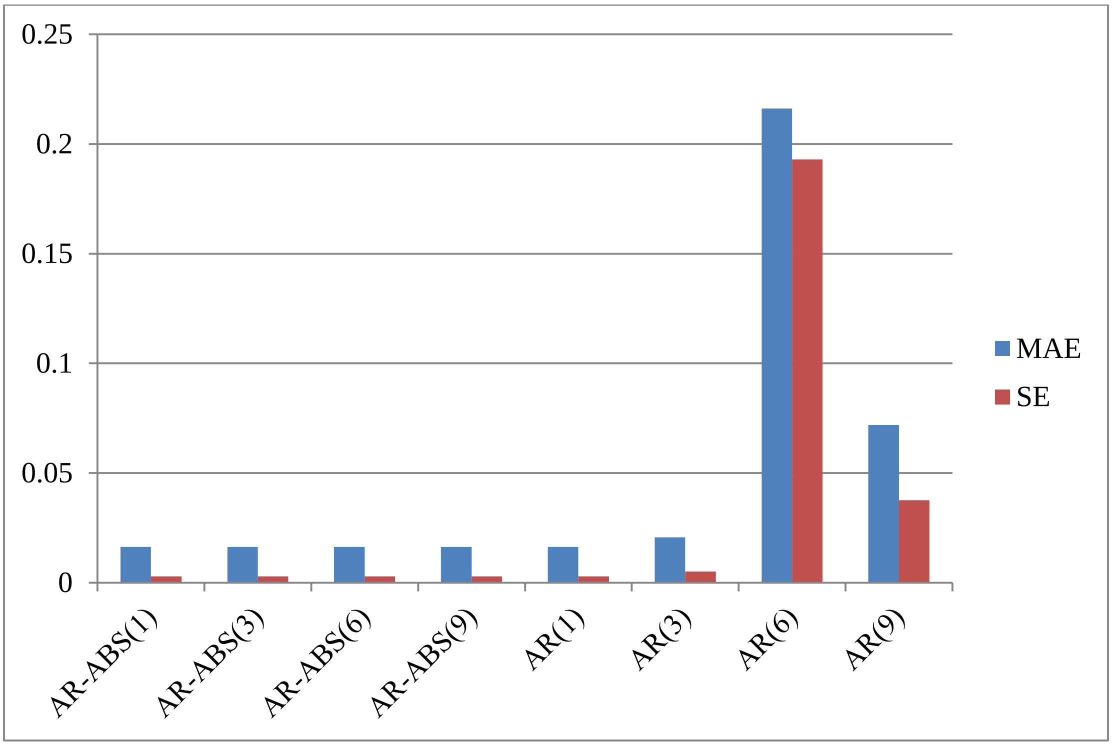

Fig. 17 *MAE and SE in real stock market, average of eight stocks, Period II, using TR6*

Here all the errors are smaller than in previous charts. The greatest error occurred by AR(6) prediction method. The chart illustrates some numerical instability of AR(9) and AR(6) model. The reason is that determinants of the corresponding system of linear equations can suddenly contract to small values due to some seemingly insignificant changes of data, for example by shifting the start data for several days. This is not a frequent event, but the possibility exists, what is illustrated by Figure 17.

The AR-ABS models based on linear programming are more stable. More stable are and AR(1) and AR(3) models because at small p the systems of equations are well defined, as usual. Therefore, using these models, the prediction errors are similar for all trading rules.

Figure 18 shows average profits obtained by eight prediction modes and

TR6.

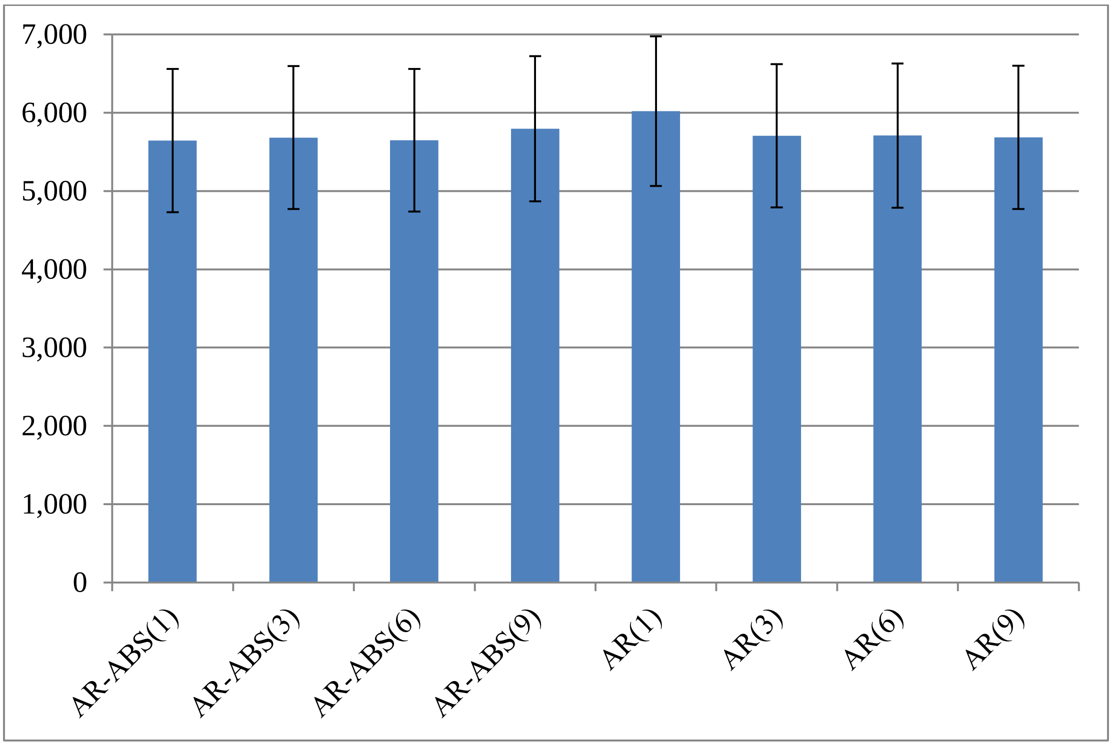

Fig. 18 *Average profits of eight prediction modes in real stock market, Period II, using TR6*

Both, the pattern and values of profits are different from TR1 shown in Figure 15. Here the profits are almost independent on prediction model and almost four times smaller comparing with the best ones of TR1.

Figure 19 shows average portfolios of eight prediction modes of TR6.

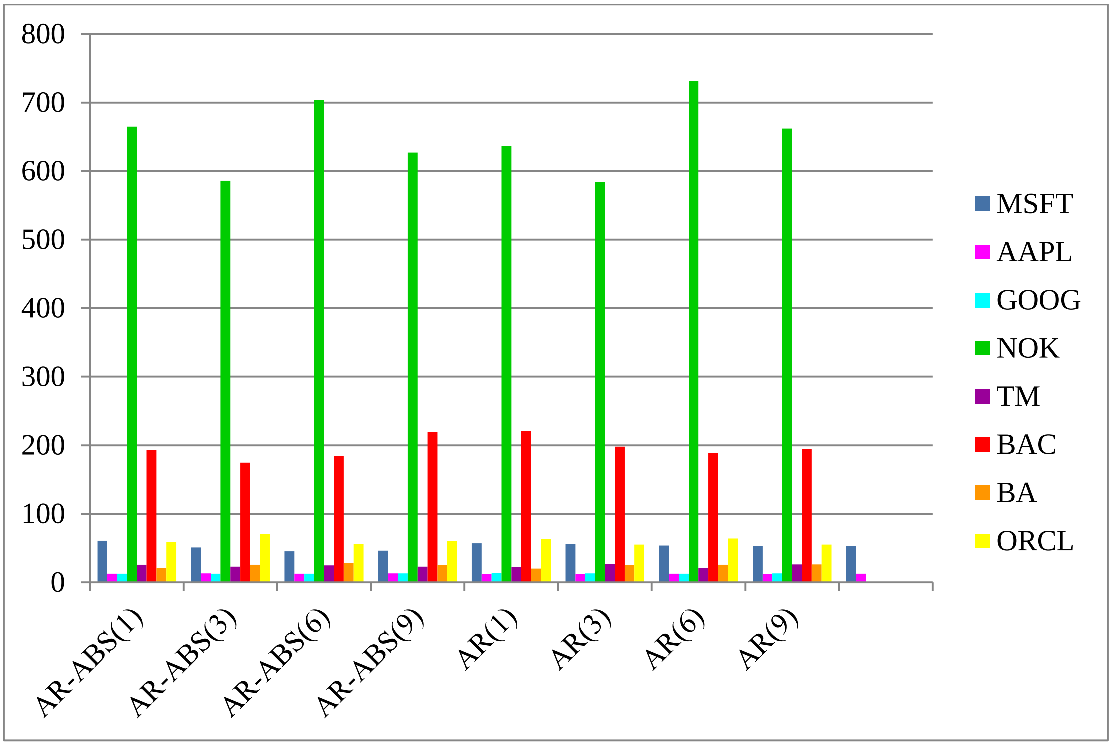

Fig. 19 *Average portfolios in real stock market, Period II, using TR6 and different prediction model*

Here all stocks are in all portfolios, but favorites are NOK and BAC stocks.

In Period II, the greatest profit was obtained using TR1 and AR(1). The average portfolio contains only BAC stocks. The explanation is that these stocks continue their recovery also in Period II and that the trading rule is not the risk-avoiding one.

Using the risk-avoiding TR6, the greatest average profit of 6019.915 was obtained using the forecast model AR(1). The diversified portfolio contains mostly NOK stocks, but also includes others: BAC, ORCL, MSFT, TM, AAPL, BA and GOOG. For comparison, the best profit of more risky TR1, which provided single-stock portfolio was 22916.86.

4.3. Real Stock Experiment – Period III

In this section, results from third period experiments are presented. Table 9 shows profits of eight prediction modes and ten trading rules. The greatest profit in this period was obtained by TR2 and AR(1). The largest losses occurred using TR9 and AR(3).

Table 9 *Average profits of eight prediction modes and ten trading rules in real stock market, Period III*

Trading Rule	AR-ABS(1)	AR-ABS(3)	AR-ABS(6)	AR-ABS(9)	AR(1)	AR(3)	AR(6)	AR(9)
TR1	126520.5	126473.7	126465	126438.2	126520.5	126478.8	126461.3	126448.4
TR2	136228.9	98385.99	120154.9	98614.86	**147125.2**	81927.59	60250.08	59100.71
TR3	49021.57	83341.16	97304.51	69018.17	59292.55	12112.26	13115.62	-4445.43
TR4	57449.19	32543.26	41944.57	39950.43	48897.3	46384.46	43837.76	36200.11
TR5	1345.915	1969.208	1496.909	1234.166	1744.026	1882.287	1636.979	1369.55
TR6	76.23677	647.0685	617.8426	553.4044	646.268	-23.4266	574.1011	433.5612
TR7	26258	11107.03	-3141.08	8858.445	27035.13	192.3318	32633.98	31632.76
TR8	10405.64	24470.35	21051.15	30343.58	20379.42	46602.13	12619.58	-2498.2
TR9	-9529.5	6425.078	-1446.04	14471.1	-9105.42	**-34398.2**	-19089.1	28801.31
TR10	18407.47	18358.39	19399.88	15135.24	17135.13	14919.72	16221.58	11851.12

The Tables 10 and 11 shows MAE and SE of eight different prediction modes.

Table 10 *MAE in real stock market, average of eight stocks, Period III*

Trading Rule	MAE							
	AR-ABS(1)	**AR-ABS(3)**	**AR-ABS(6)**	**AR-ABS(9)**	**AR(1)**	**AR(3)**	**AR(6)**	**AR(9)**
TR1	0.0145465	0.0145207	0.014607	0.0145859	0.0145369	0.018539	0.026759	0.040131
TR2	0.0143822	0.0144287	0.014386	0.01443	0.0144061	0.01914	0.171801	0.374929
TR3	0.014421	0.0144568	0.014427	0.0144854	0.0144389	0.019158	0.171687	0.375372
TR4	0.0143073	0.0143458	0.014323	0.014321	0.0143352	0.01613	0.019211	0.023536
TR5	0.0142862	0.0142967	0.014394	0.0144153	0.0143063	0.017005	0.027053	0.030804
TR6	0.0142871	0.0143277	0.01435	0.0143714	0.0143327	0.017124	0.020764	0.026336
TR7	0.0143182	0.0142837	0.014337	0.0143977	0.0143485	0.021629	0.05214	0.027531
TR8	0.0142675	0.0142643	0.014339	0.0143702	0.0143066	0.015963	0.023303	0.026653
TR9	0.014259	0.0143022	0.014366	0.014424	0.0142877	0.018033	0.08783	0.025713

TR10	0.014259	0.0143022	0.014366	0.014424	0.0142877	0.018033	0.08783	0.025713

Table 11 *SE in real stock market, average of eight stocks, Period III*

Trading Rule	**SE**							
	AR-ABS(1)	**AR-ABS(3)**	**AR-ABS(6)**	**AR-ABS(9)**	**AR(1)**	**AR(3)**	**AR(6)**	**AR(9)**
TR1	0.002425	0.0024244	0.002426	0.0024249	0.0024233	0.003672	0.007888	0.012006
TR2	0.002369	0.0023698	0.002369	0.0023685	0.0023681	0.004674	0.150096	0.336021
TR3	0.0023741	0.0023749	0.002375	0.0023759	0.002373	0.004682	0.150084	0.336627
TR4	0.0023357	0.0023374	0.002337	0.0023378	0.0023352	0.002764	0.003445	0.004369
TR5	0.0023964	0.002397	0.002401	0.0023996	0.0023967	0.003124	0.009095	0.009403
TR6	0.0024009	0.0024021	0.002403	0.0024037	0.0024015	0.003205	0.004055	0.00601
TR7	0.002404	0.0024035	0.002405	0.0024067	0.0024038	0.007094	0.03373	0.006087
TR8	0.0023859	0.0023856	0.002388	0.0023887	0.0023855	0.002941	0.006044	0.006917
TR9	0.0024492	0.0024507	0.002452	0.0024527	0.0024496	0.004008	0.069795	0.005111
TR10	0.0024492	0.0024507	0.002452	0.0024527	0.0024496	0.004008	0.069795	0.005111

Here, the numerical instability of AR(6) and AR(9) models is illustrated by greater than usual prediction errors obtained using trading rules No. 2 and No. 3. Using other trading rules, the patterns of errors are similar to earlier periods, what is illustrated by Figure 20.

Table 12 *Average portfolios of ten trading rules in real stock market, Period III*

Stock name	**TR1**	**TR2**	**TR3**	**TR4**	**TR5**
	AR-ABS1	**AR-ABS3**	**AR-ABS3**	**AR-ABS6**	**AR9**
MSFT	0	0	0	17.1342	104.7107
AAPL	0	0	7.9205	5	3.8195
GOOG	0	0	5.9534	22.3178	5.9942
NOK	53901.9452	58318.7836	30566.9041	18627.8027	1285.9027
TM	0	0	14.4082	39.6247	29.8233
BAC	0	0	3565.4849	3819.6548	229.2066
BA	0	0	99.1288	13.1014	35.026
ORCL	0	0	110.9671	61.8082	91.0553
AVG (strategy)	**6737.74**	**7289.85**	**4296.35**	**2825.81**	**223.19**
Stock name	**TR6**	**TR7**	**TR8**	**TR9**	**TR10**
	AR-ABS1	**AR-ABS3**	**AR-ABS3**	**AR-ABS6**	**AR9**
MSFT	58.9784	0	0.0164	59.7288	74.8438
AAPL	12.2425	0	5.6438	6.937	7.0411
GOOG	11.451	0	10.4795	2.4055	19.937

NOK	419.9153	20705.5644	2503.7425	1977.6192	616.8986
TM	24.9707	0	3.9452	345.1342	3.3726
BAC	155.5866	0	4257.3589	570.1096	2624.7836
BA	25.1384	0	19.0137	14.7151	27.8137
ORCL	39.3485	0	10.0877	32.7068	52.8137
AVG (strategy)	**93.45**	**2588.20**	**851.29**	**376.17**	**428.44**

Here most of portfolios include NOK stocks. Some portfolios include other stocks, too: BAC, AAPL, MSFT.

In this period, average profits of all prediction models are almost the same, while the prediction errors differ, see Figure 20. This is an additional illustration of unexpected relation between the profits and prediction accuracy, see Figures 34-37.

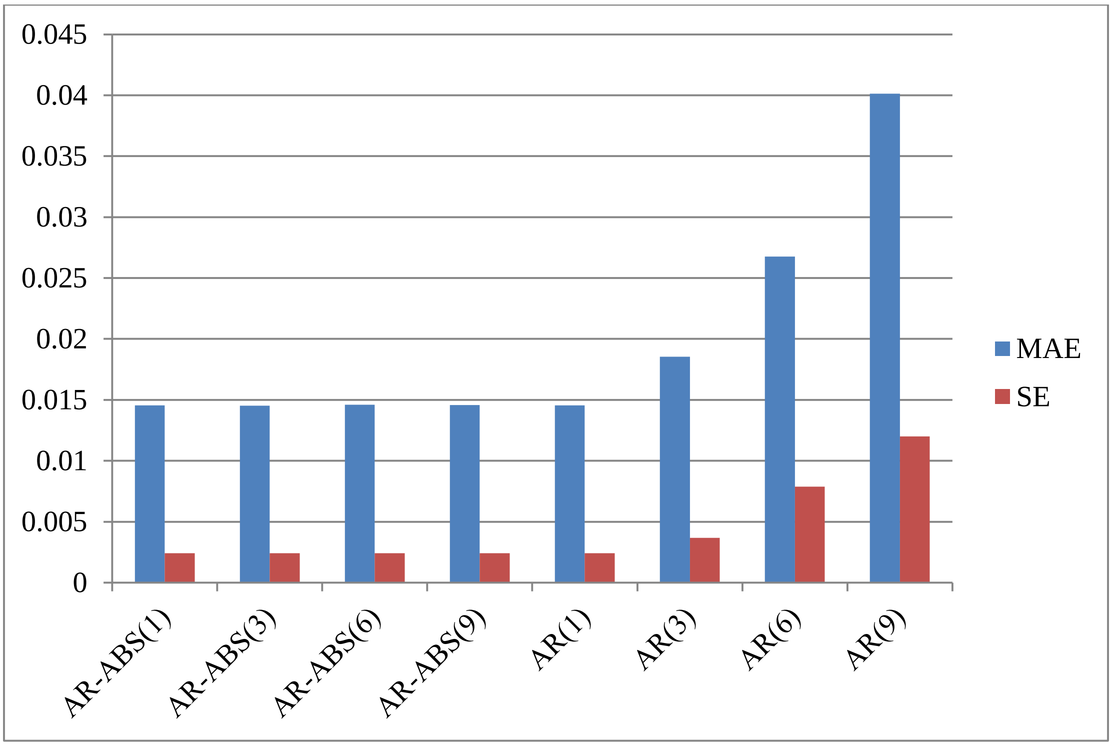

Fig. 20 MAE and SE in real stock market, average of eight stocks, Period III, using TR1

Figure 21 illustrates relatively stable present economic conditions. The only exception is the jump of Nokia (NOK) stocks.

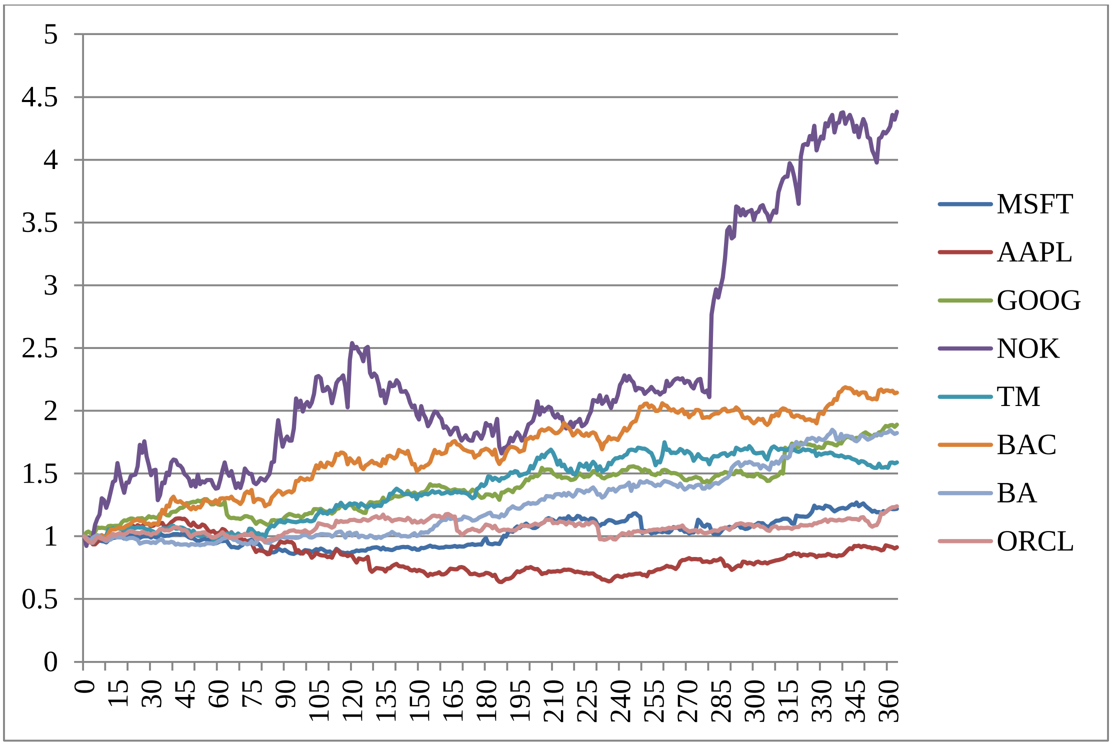

Fig. 21 *Normalized daily prices of eight stocks in Period III*

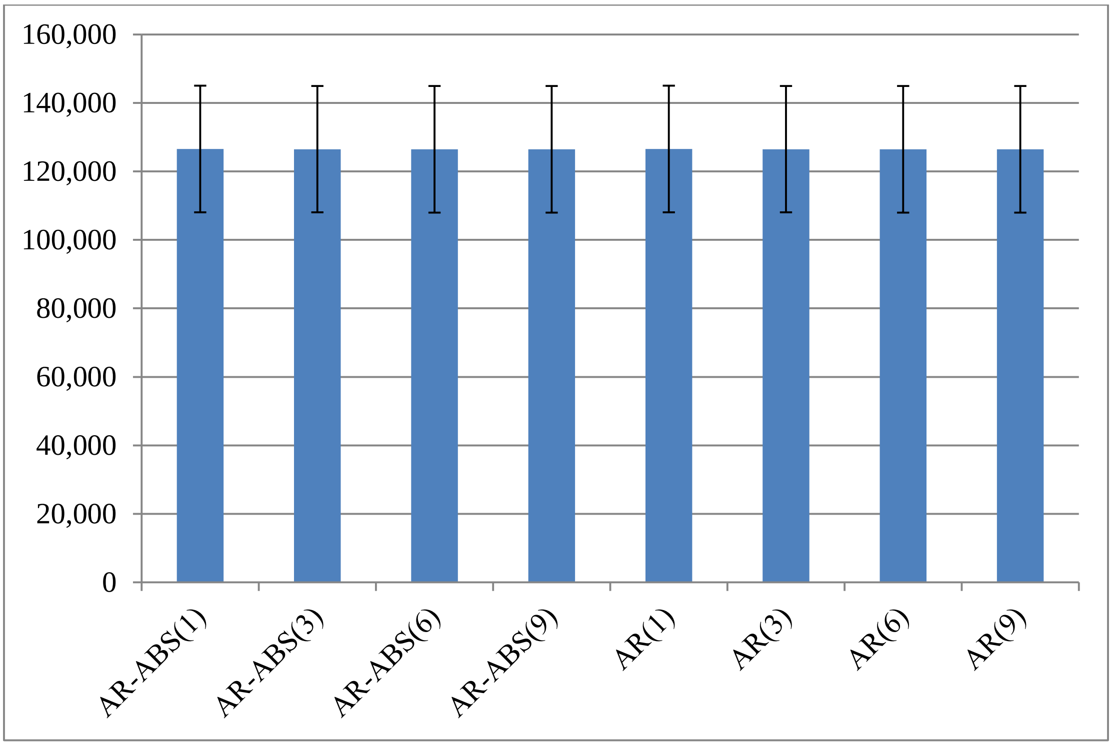

Fig. 22 *Average profits of eight prediction modes in real stock market, Period III, using TR1*

Figure 22 shows average profits obtained using eight prediction modes and TR1.

Figure 23 shows average portfolios by eight prediction modes, using TR1.

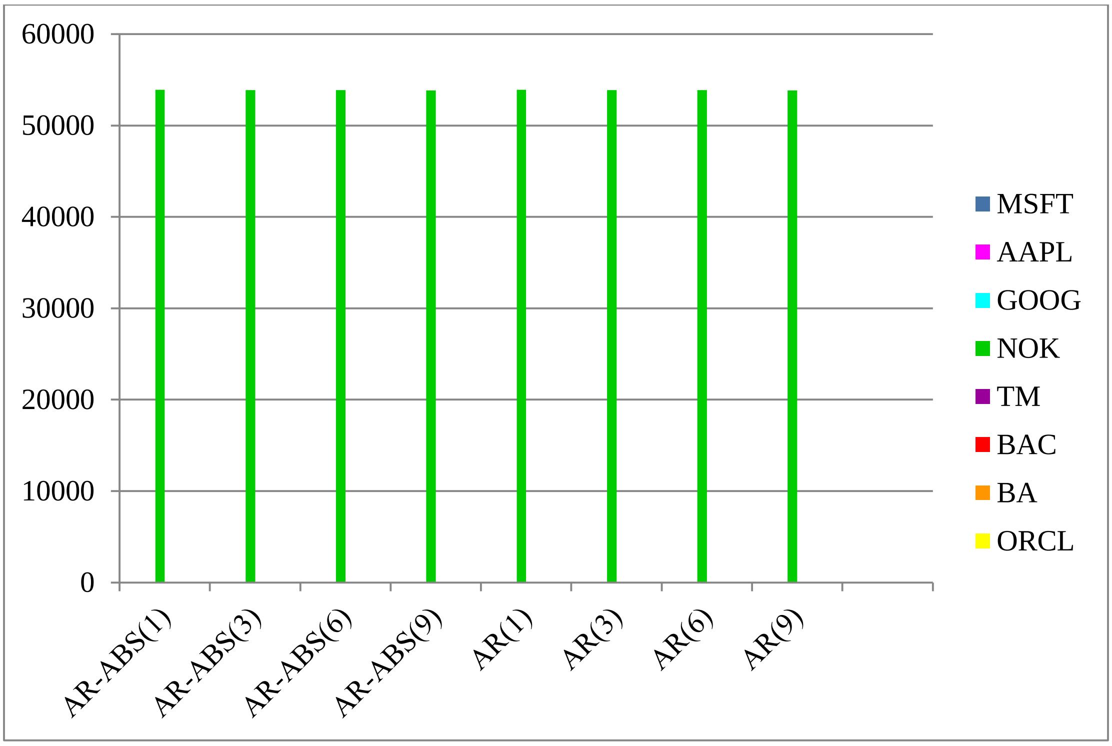

Fig. 23 *Average portfolios in real stock market, Period III, using TR1 and different prediction modes*

In Figure 23 all the portfolios include only NOK stocks, because in this period, the NOK stocks were recovering after the previous losses. The greatest profit was obtained using TR2 and AR-ABS(1). The best average portfolio contains only NOK stocks.

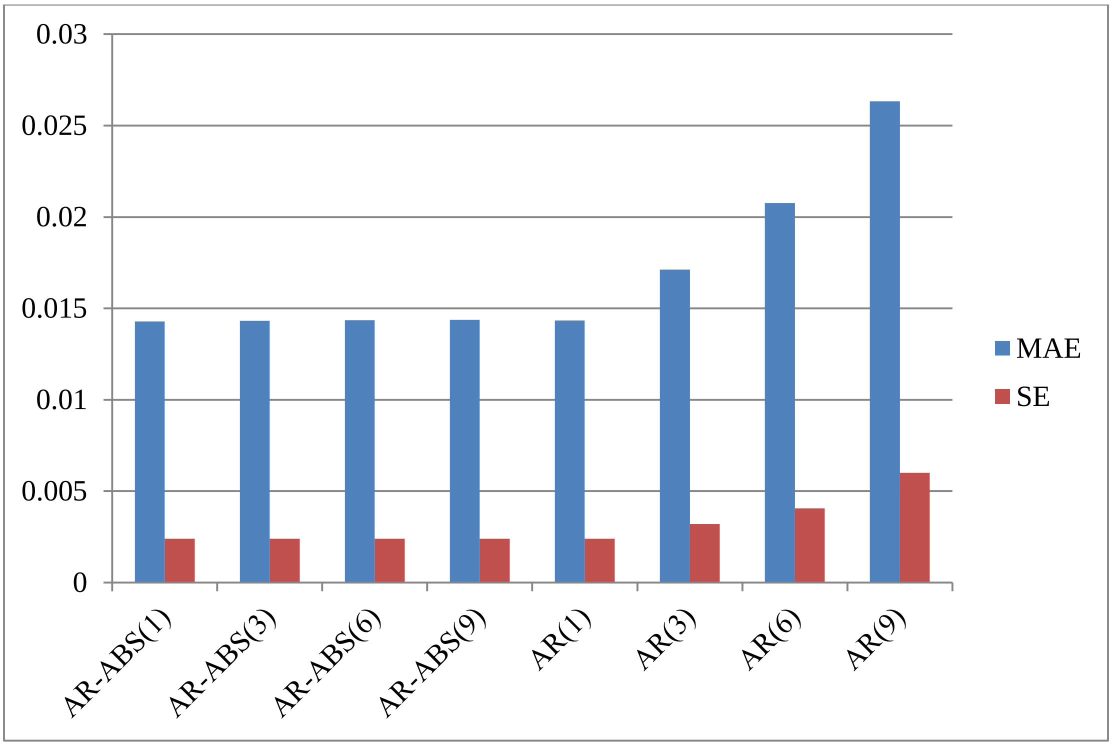

Fig. 24 *MAE and SE in real stock market, average of eight stocks, Period III, using TR6*

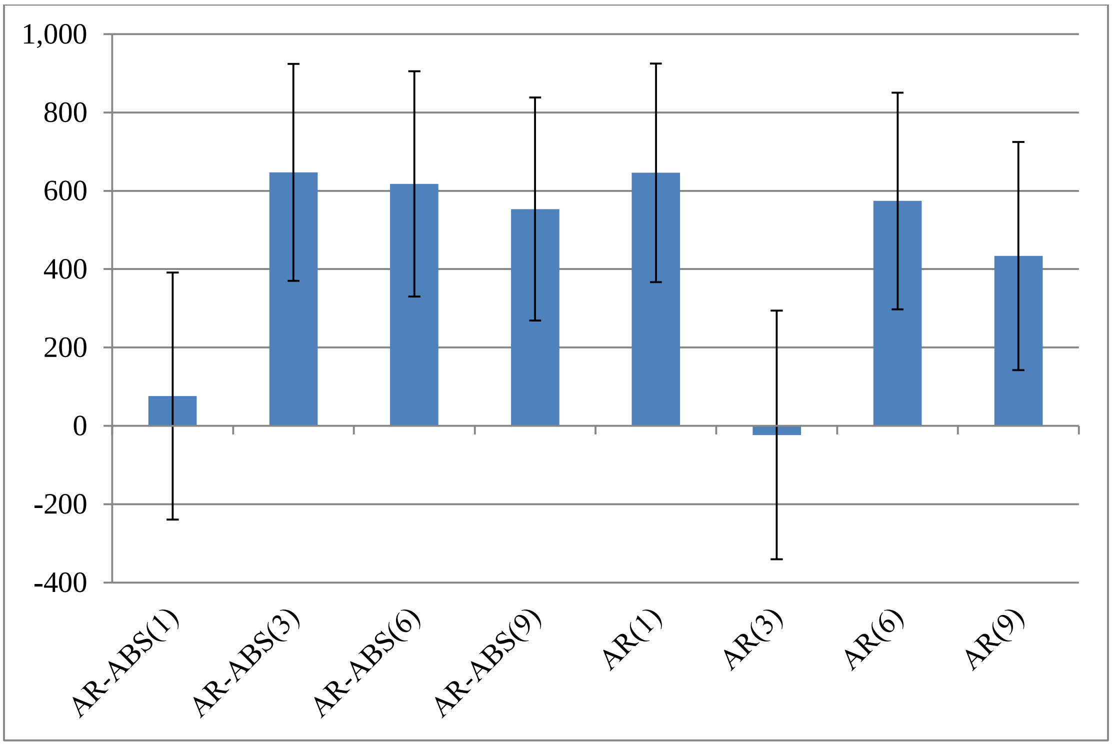

Fig. 25 *Average profits of eight prediction modes in real stock market, Period III, using TR6*

Figure 24 presents average forecast errors of 8 traders using different forecast methods by TR6.

Figure 25 shows profits of TR6 and eight prediction modes. In Figure 25 profits are different as compared to other periods and other trading rules. The greatest profit is obtained by prediction model AR(1) closely followed by AR-ABS(3).

Figure 26 shows average portfolios.

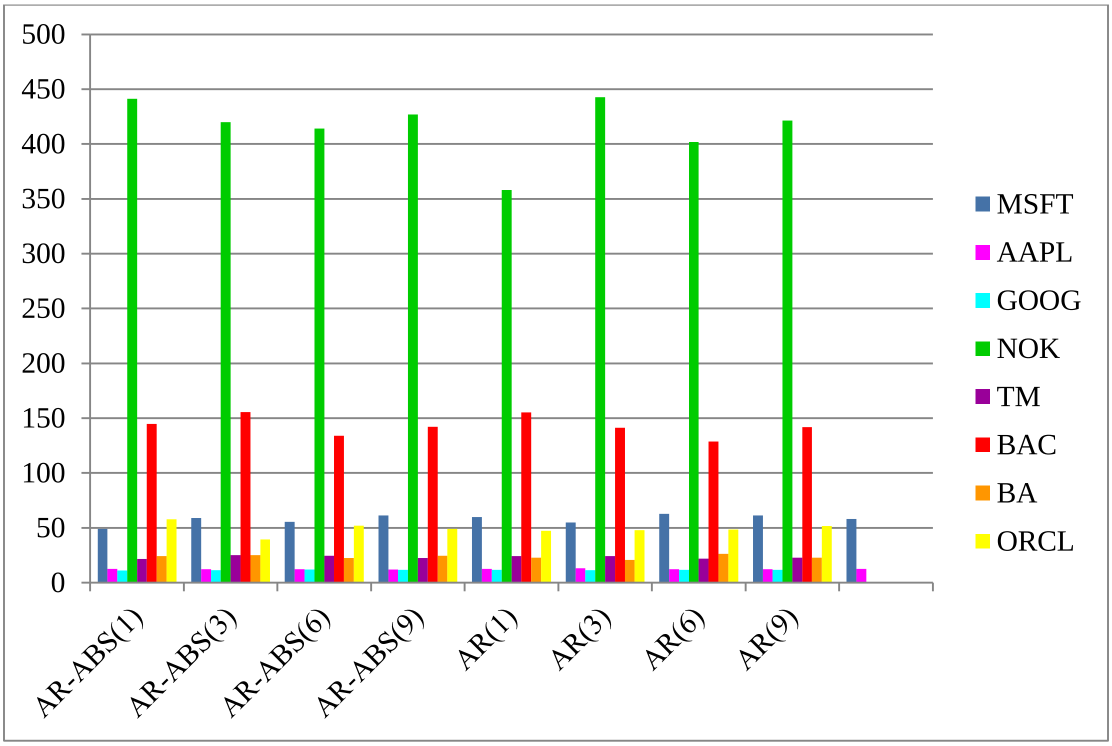

Fig. 26 *Average portfolios in real stock market, Period III, using TR6 and different prediction modes*

All the portfolios prefer NOK and BAC stocks. They include all other stocks too, but in lesser proportions.

In this period, using risk-avoiding TR6, the greatest profit was obtained by the forecast model AR-ABS(3). The average portfolio contains mostly NOK stocks. It includes all others: BAC, MSFT, ORCL, TM, BA, AAPL and GOOG, but in lesser proportions. The profit of the portfolio 617.84 was very small as compared with the profit 147125.2 of the single-stock portfolio provided by more risky TR2.

4.4. Virtual Stock Experiment

In this section, the experiments using virtual data and four short-term trading rules are discussed. Table 13 shows average profits of eight prediction strategies using four trading rules. Here the greatest profit was obtained by TR3 and AR(9) strategy and the biggest losses occurred using TR1 and AR-ABS(3).

Table 13 *Average profits of eight prediction modes and four trading rules in virtual stock market*

Trading Rule	AR-ABS(1)	AR-ABS(3)	AR-ABS(6)	AR-ABS(9)	AR(1)	AR(3)	AR(6)	AR(9)
TR1	3381.62	**-3024.88**	17577.00	-2795.87	739.39	1347.49	13689.31	833.36
TR2	-29.0392	901.9031	13596.29	2433.077	-388.537	1010.475	435.0855	4590.43
TR3	-290.635	16527.23	4541.016	5683.213	-402.586	8273.673	39088.07	**61273.55**
TR4	-182.831	9262.315	3262.828	639.9958	-311.727	-404.884	-202.49	-198.736

Tables 14 and 15 show prediction errors. The largest errors occurred by AR(6) and AR(9). It shows that in the virtual environment, simple models provide lesser errors, too.

Table 14 *MAE in virtual stock market, average of eight stocks*

Trading Rule	MAE							
	AR-ABS(1)	AR-ABS(3)	AR-ABS(6)	AR-ABS(9)	AR(1)	AR(3)	AR(6)	AR(9)
TR1	0.011305	0.011293	0.011477	0.011386	0.0111	0.013842	0.034685	0.044407
TR2	0.011043	0.011125	0.011085	0.011019	0.010848	0.012841	0.023406	0.096941
TR3	0.009789	0.009948	0.010182	0.010302	0.009966	0.011636	0.033677	0.079483
TR4	0.009816	0.010176	0.010345	0.010436	0.009784	0.011677	0.038337	0.124938

Table 15 *SE in virtual stock market, average of eight stocks*

Trading Rule	SE							
	AR-ABS(1)	AR-ABS(3)	AR-ABS(6)	AR-ABS(9)	AR(1)	AR(3)	AR(6)	AR(9)

TR1	0.002523	0.00251	0.002516	0.002514	0.002513	0.003434	0.011968	0.013409
TR2	0.002548	0.002538	0.002538	0.002537	0.00254	0.002866	0.006353	0.041476
TR3	0.002445	0.002438	0.002444	0.002445	0.002449	0.002646	0.013495	0.041892
TR4	0.002469	0.002469	0.002473	0.002476	0.00247	0.003024	0.016161	0.086981

Table 16 shows the portfolios of eight prediction modes and four trading rules. Here all stocks are included in all portfolios, but most popular are the second, sixth and eighth stocks.

Table 16 *Average portfolios of four trading rules in virtual stock market*

	TR1	**TR2**	**TR3**	**TR4**
	AR-ABS(6)	**AR-ABS(6)**	**AR(9)**	**AR-ABS(3)**
first	256.20	81.97	208.43	397.70
second	1888.81	471.39	1518.32	663.39
third	23.01	72.59	57.41	58.31
fourth	1.99	83.97	47.24	629.41
fifth	97.83	629.77	313.46	481.58
sixth	564.96	47.75	90.74	8.53
seventh	125.32	171.09	68.24	493.38
eighth	1337.56	540.44	1326.49	1237.99
AVG (strategy)	**536.96**	**262.37**	**453.79**	**496.29**

Figure 27 shows average prediction errors of eight prediction modes, using TR1. The pattern of errors is similar to real stocks environment. Numerical values are not very different, too.

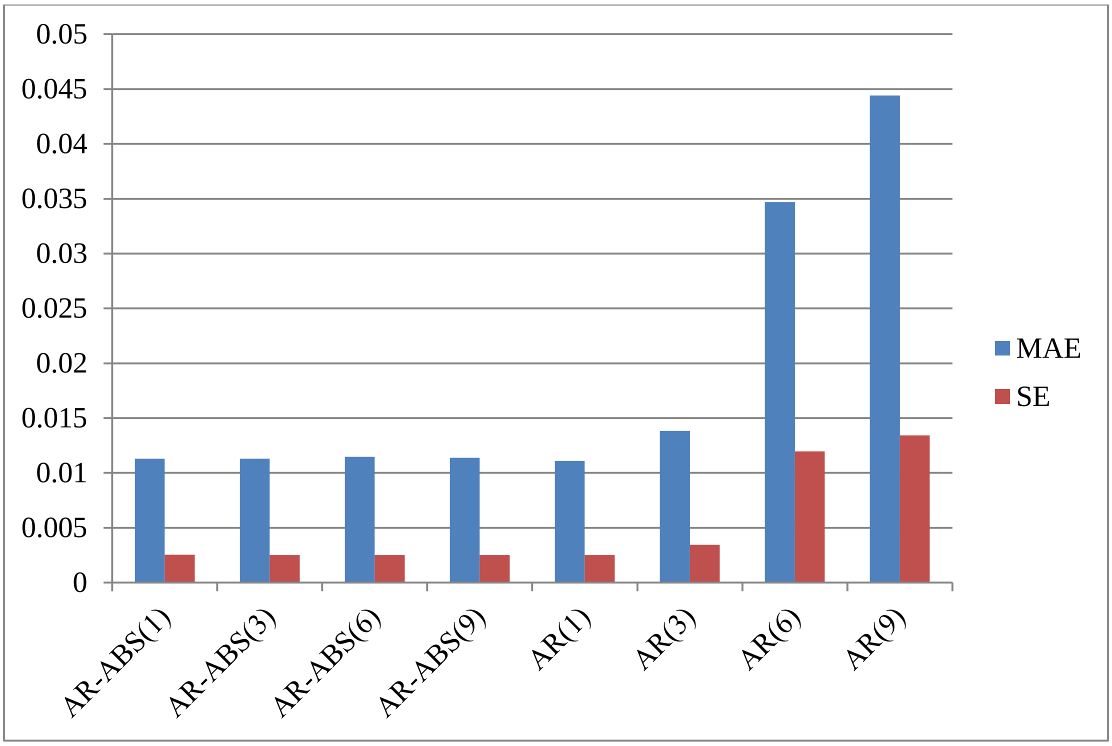

Fig. 27 *MAE and SE in virtual stock market, average of eight stocks, using TR1*

Figure 28 shows the normalized stock prices in virtual stock market.

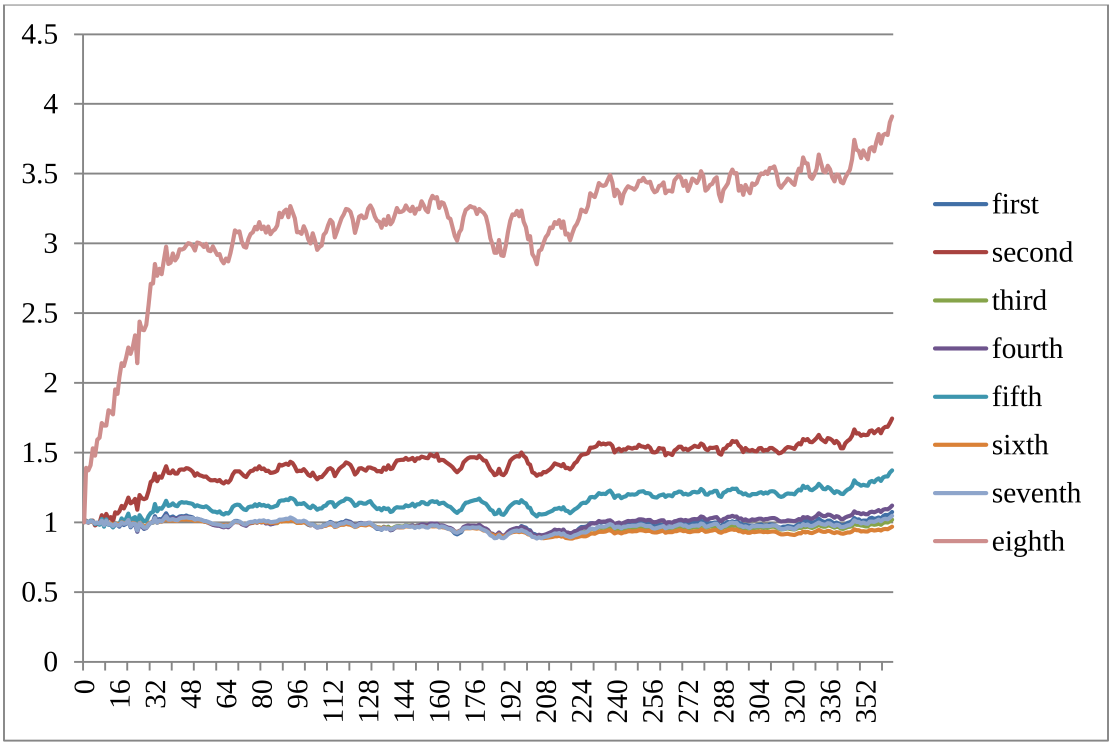

Fig. 28 *Normalized average daily prices of eight different virtual stocks*

Figure 29 shows average profits obtained by eight prediction modes, using TR1. The greatest profit was obtained by AR-ABS(6), the largest losses happened using AR-ABS(3). Note that prediction errors of these two prediction models are almost identical, see Figure 27.

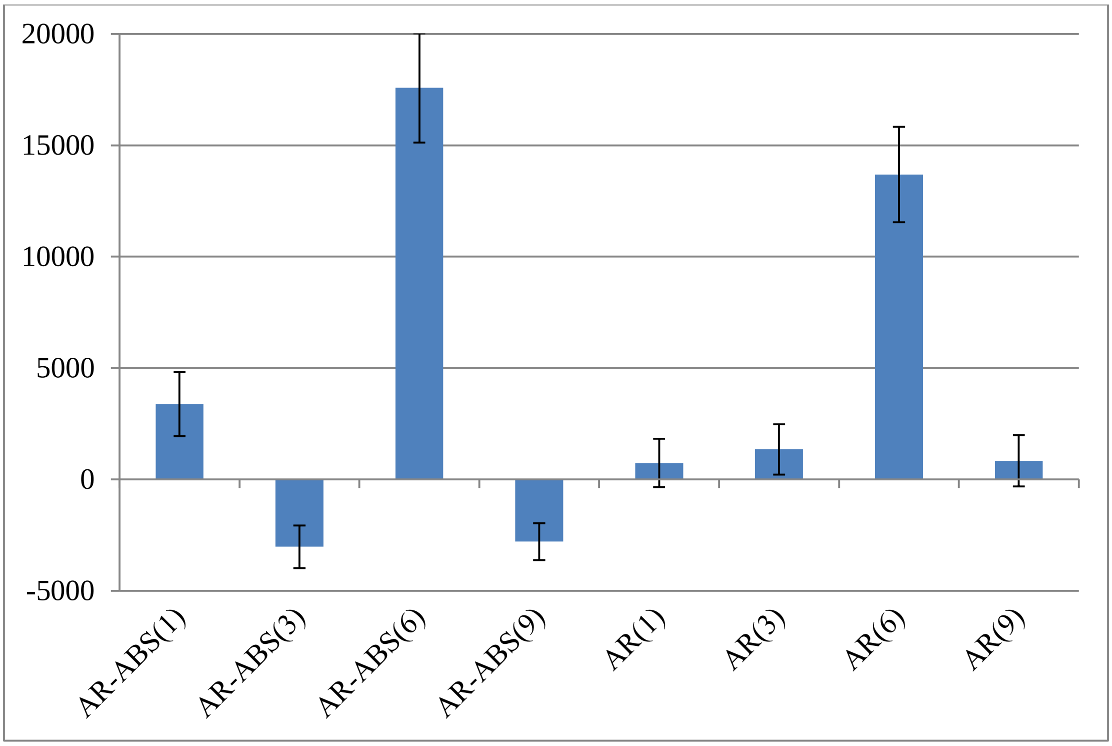

Fig. 29 *Average profits of eight prediction modes in virtual stock market, using TR1*

Figure 30 shows average portfolios of eight prediction modes, using TR1. Using this trading, the most profitable portfolio obtained by AR-ABS(6) includes a mixture of eight different.

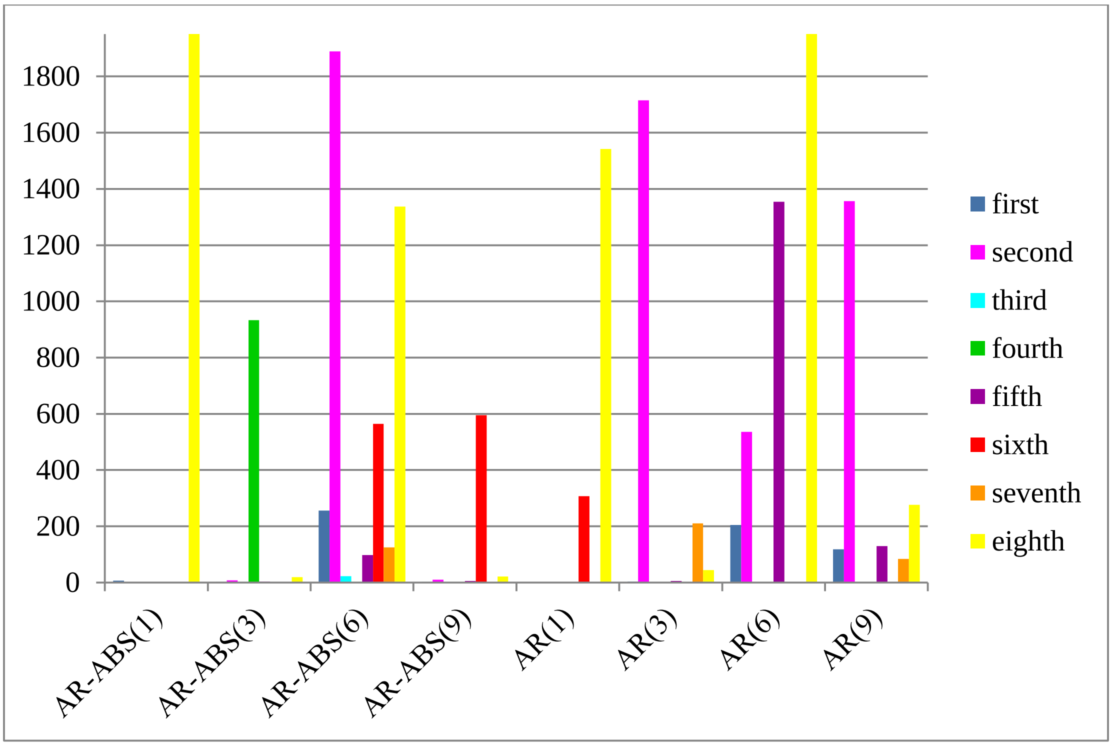

Fig. 30 *Average portfolios in virtual stock market, using TR1 and different prediction modes*

The greatest profit was obtained using the forecast model AR-ABS(6) and the average portfolio consists mostly from the fifth stock, but also includes six others.

For comparison, Figure 31 shows average prediction errors of eight prediction modes, using other trading rule – TR4. The pattern is similar, but errors are lesser in comparison with TR1.

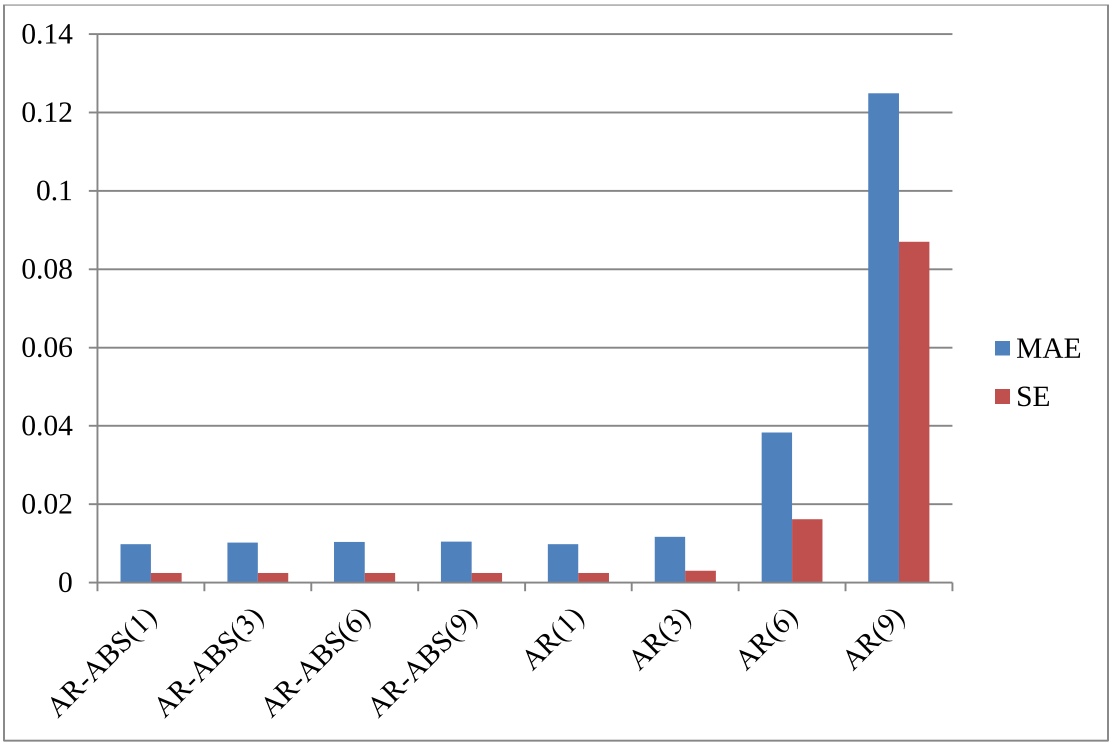

Fig. 31 *MAE and SE in virtual stock market, average of eight stocks, using TR4*

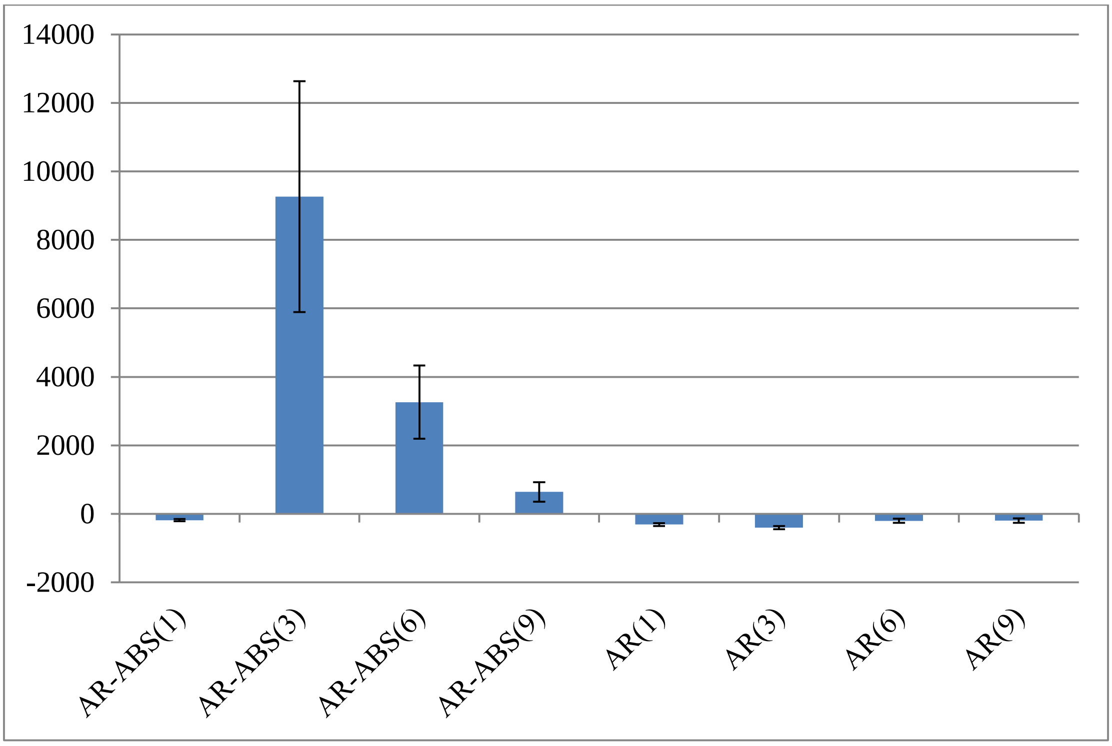

Fig. 32 *Average profits of eight prediction modes in virtual stock market, using TR4*

Figure 32 shows the corresponding average profits obtained using TR4. The greatest profit was achieved by AR-ABS(3). Note, that using AR-ABS(1) with the same prediction error the losses occurred, instead of profits, what illustrates the complicated relation of profits to prediction accuracy also in the virtual environment.

Figure 33 shows average portfolios of eight prediction modes, using TR4. The most profitable portfolio obtained by AR-ABS(3) includes a mixture of all stocks.

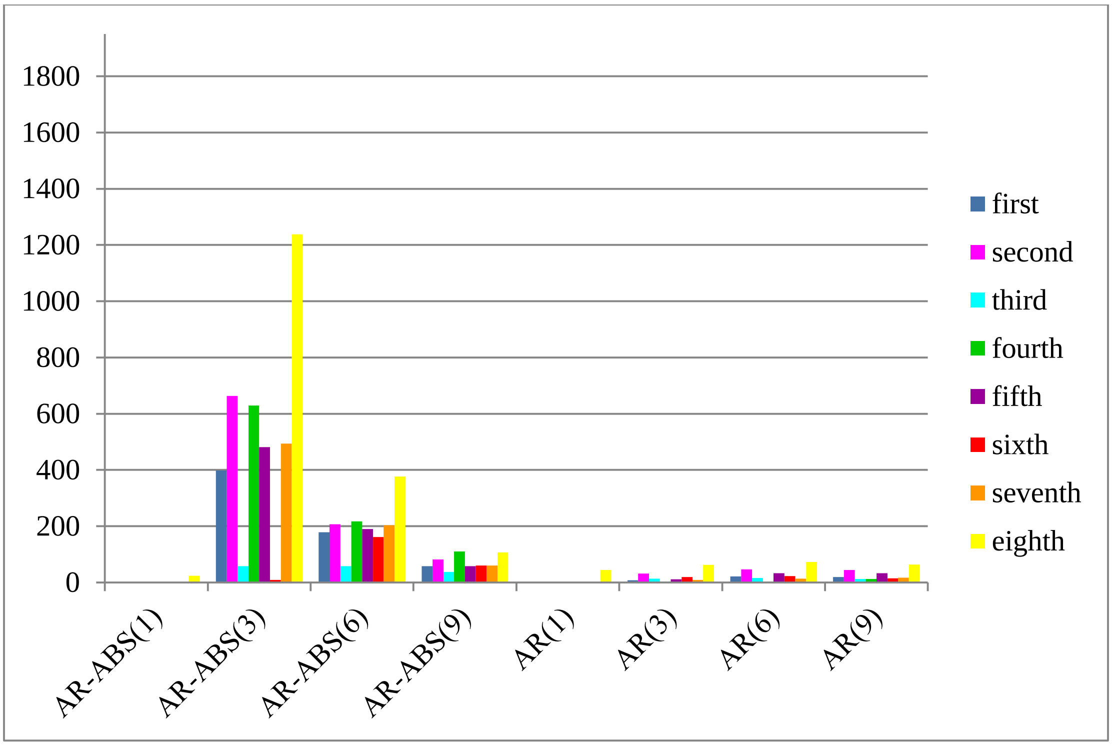

Fig. 33 *Average portfolios in virtual stock market, using TR4 and different prediction modes*

Using the risk-averse TR4, the greatest profit was obtained by the forecast model AR-ABS(3). It contains almost all stocks, only sixth stock was ignored. Here the best profit 9262.37 is about two times less than the profit 17577.00 provided by the more risky TR1.

4.5. On the Correlation Between the Prediction Errors and Actual Profits

Comparing the figures showing prediction errors in MAE and SE with the figures representing the average profits, we see that minimal MAE and SE do not necessarily provide the maximal profits. This contradicts the general opinion that the investors which predict stock prices better are rewarded by higher profits. To illustrate this paradoxical situation further we show the correlations between the average prediction errors and average profits.

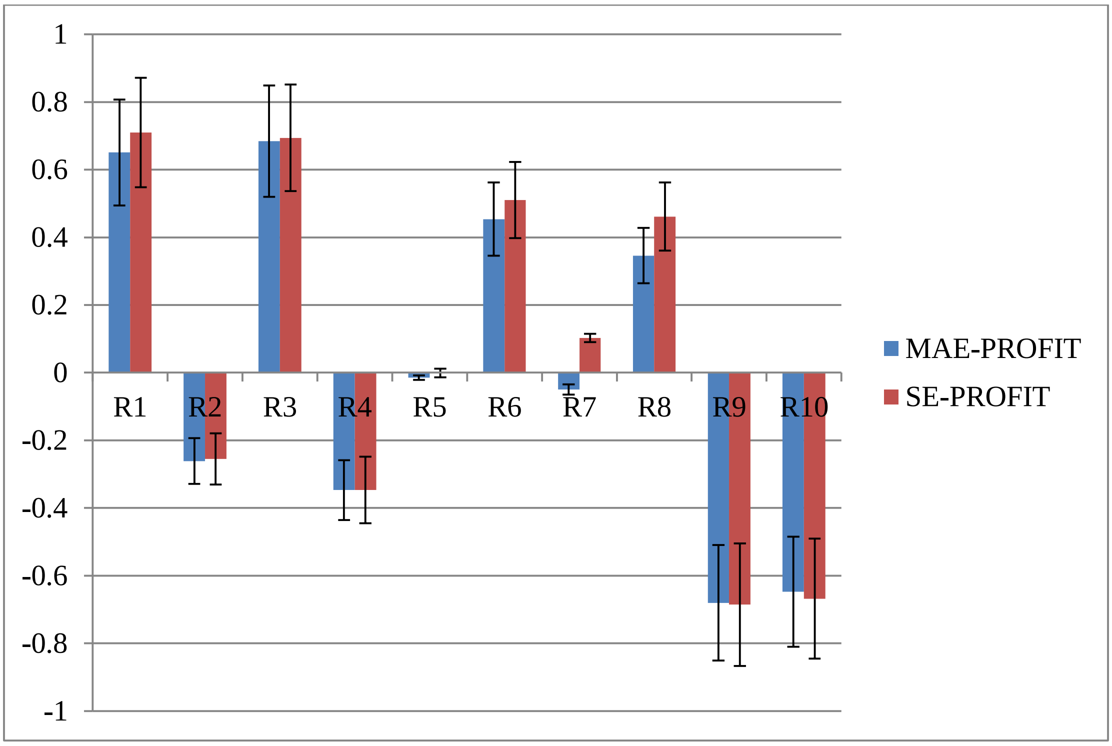

Fig. 34 *Correlation of profits and prediction errors in Period I*

Figure 34 shows the correlation of actual profits and prediction errors during the post-crisis recovery time. Contrary to reasonable expectations, the correlation is positive for four MAE and five SE of ten trading rules. This means that prediction models with larger errors provides greater profits in half of cases. 95% confidence intervals show that the differences between the correlation coefficients are not random. To explain this contradiction further investigation is planned.

Figure 35 shows the correlation of actual profits and prediction errors during more stable time. As expected, in most of trading rules, the correlation is negative, meaning that prediction models with smaller prediction errors provide greater profits. However, there is one exception: using TR5 the correlation is positive.

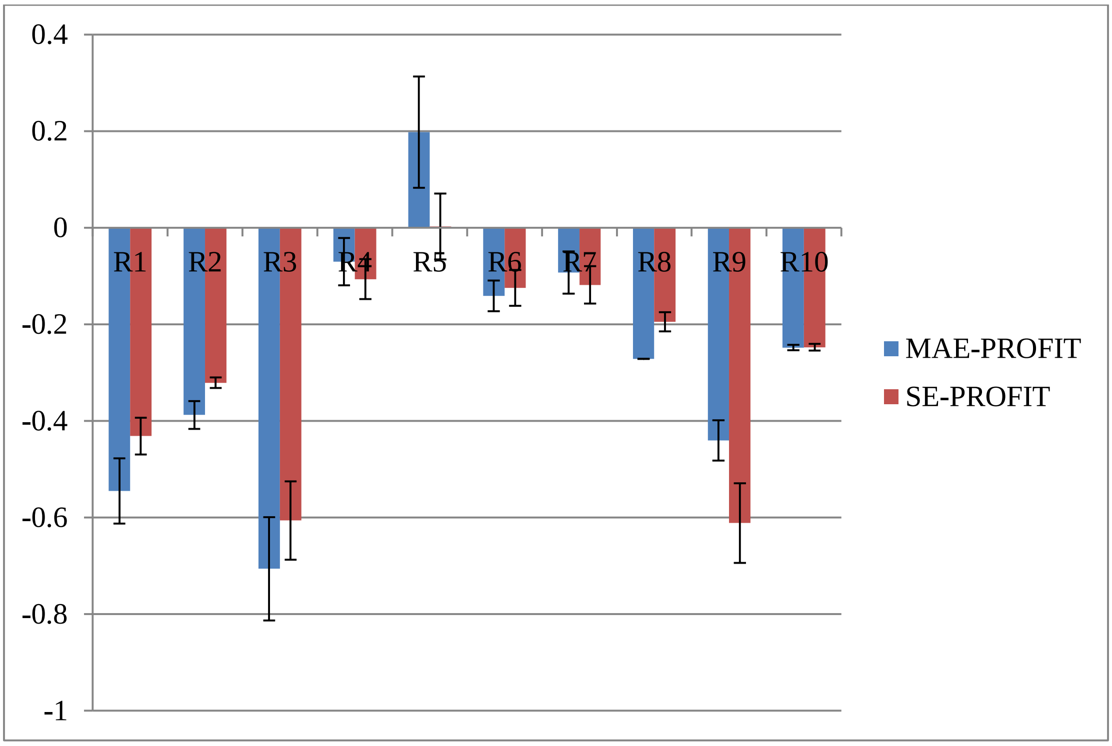

Fig. 35 *Correlation of profits and prediction errors in Period II*

Figure 36 shows the correlation of actual profits and prediction errors during the recent times. In most of trading rules, the correlation is negative. However, using TR7 the correlation is positive. The correlation is weak in all the cases.

Figure 37 shows the correlation of profits and prediction errors in the virtual market. Using three of four trading rules, the correlation is positive. The results are close to the recovery period of real market. However, using strategy R3 in the virtual market the correlation is positive and close to 1. Additional experiments are paned to explain this. Comparing the stock price graphs, we see the considerable growth in both the cases. This is a possible explanation.

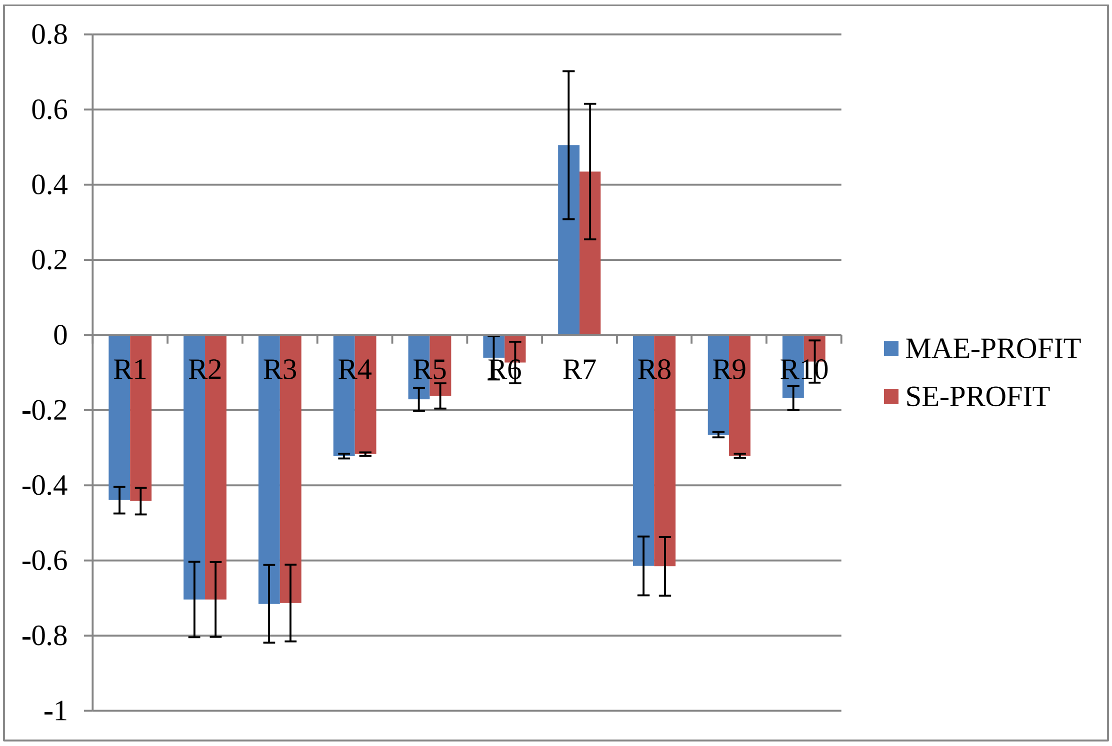

Fig. 36 *Correlation of profits and prediction errors in Period III*

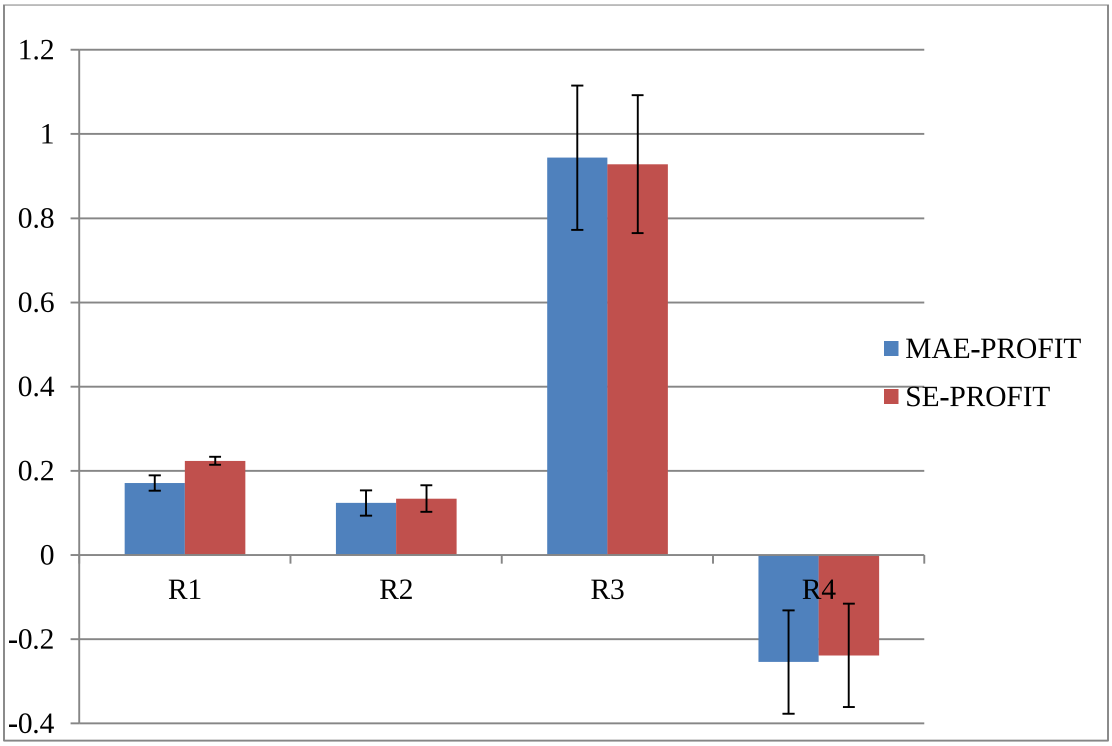

Fig. 37 *Correlation of profits and prediction errors in virtual stock market*

4.6. Investigation of Random Walk (RW)

The Random Walk (RW) strategy follows from the efficient market theory (Fama, 1995). This theory asserts that the market price reflects the real value of assets, so the best prediction strategy is the random walk. In the short trading rules, the mathematical representation of RW is the Wiener model. Therefore, an additional investigation was performed with the aim to compare statistical errors and profits of the Wiener model with autoregressive models using different investment strategies.

In the longer term strategies Wiener model behaves differently by selecting the best portfolio using the average results of all the learning period, so deviating from the basic assumption of the efficient market theory that the asset prices at the given time represents they real value.

Two short time trading rules (No. 1 and No. 4) and two long time trading rules (No. 5 and No. 6) were investigated. The data was the shorter time series (of 180 working days instead of usual 360) recorded between the second and third periods of time.

Figure 38 shows average profits of TR1. We see that the virtual market profits are almost independent on the memory length p in both the AR and AR-ABS models.

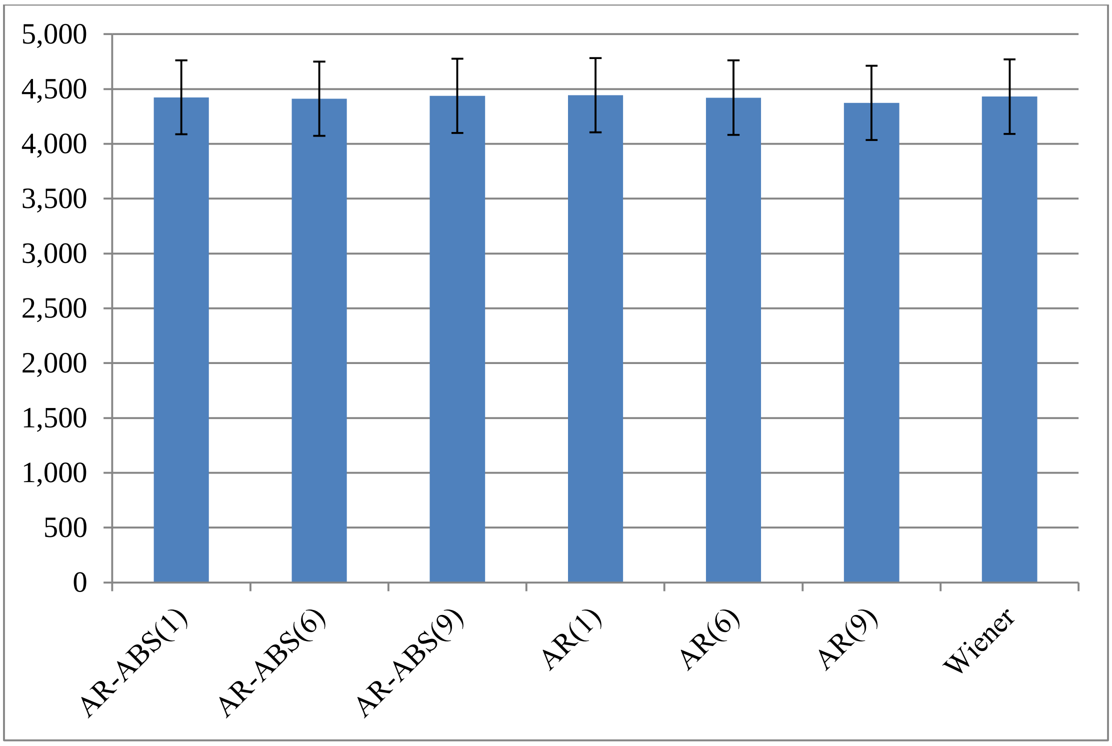

Fig. 38 *Average profits of TR1*

Figure 39 shows daily profits of TR1.

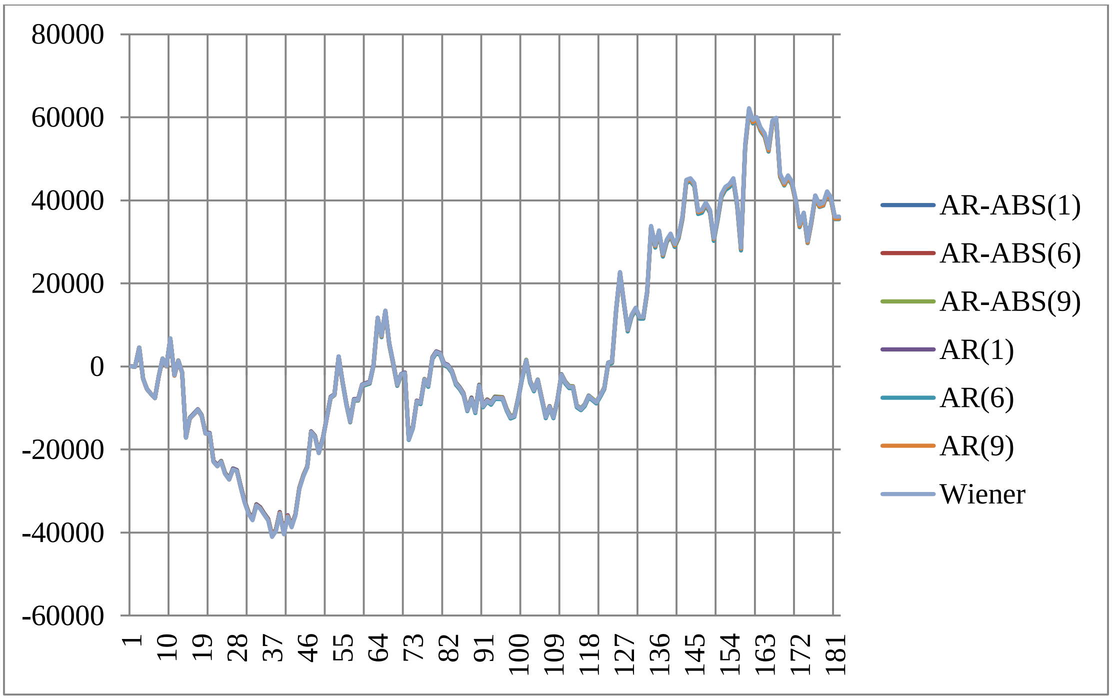

Fig. 39 *Daily profits of TR1*

Figure 40 shows average profits of TR4.

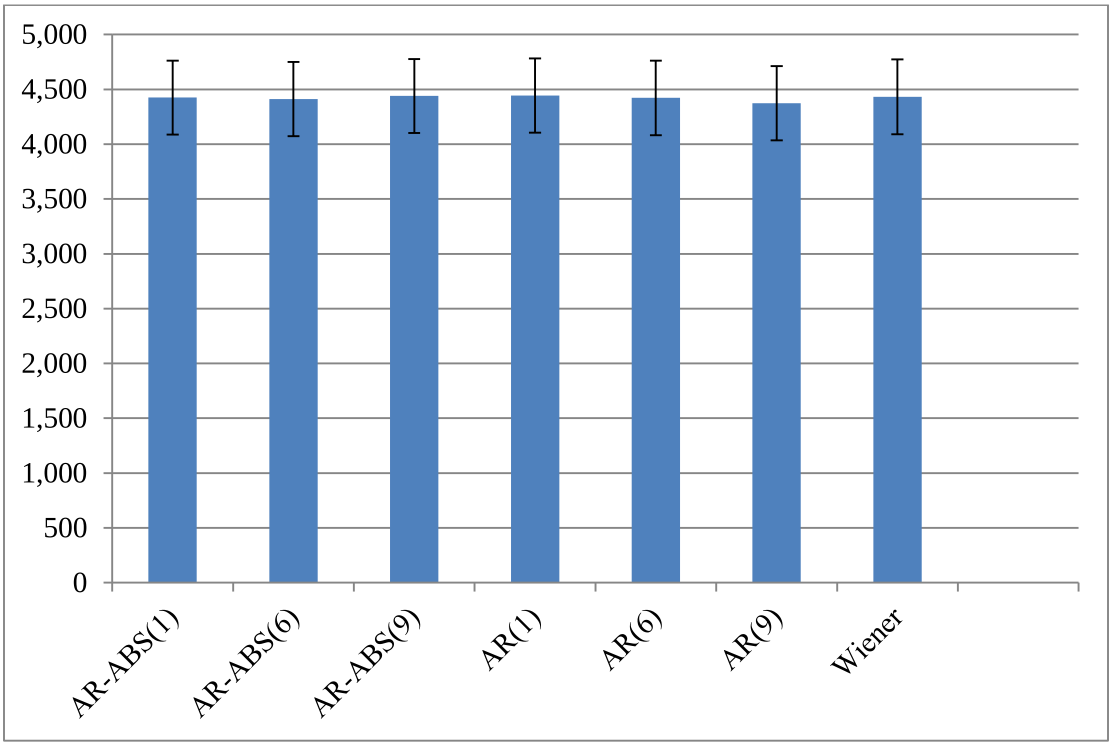

Fig. 40 *Average profits of TR4*

Figure 41 shows daily profits of TR4.

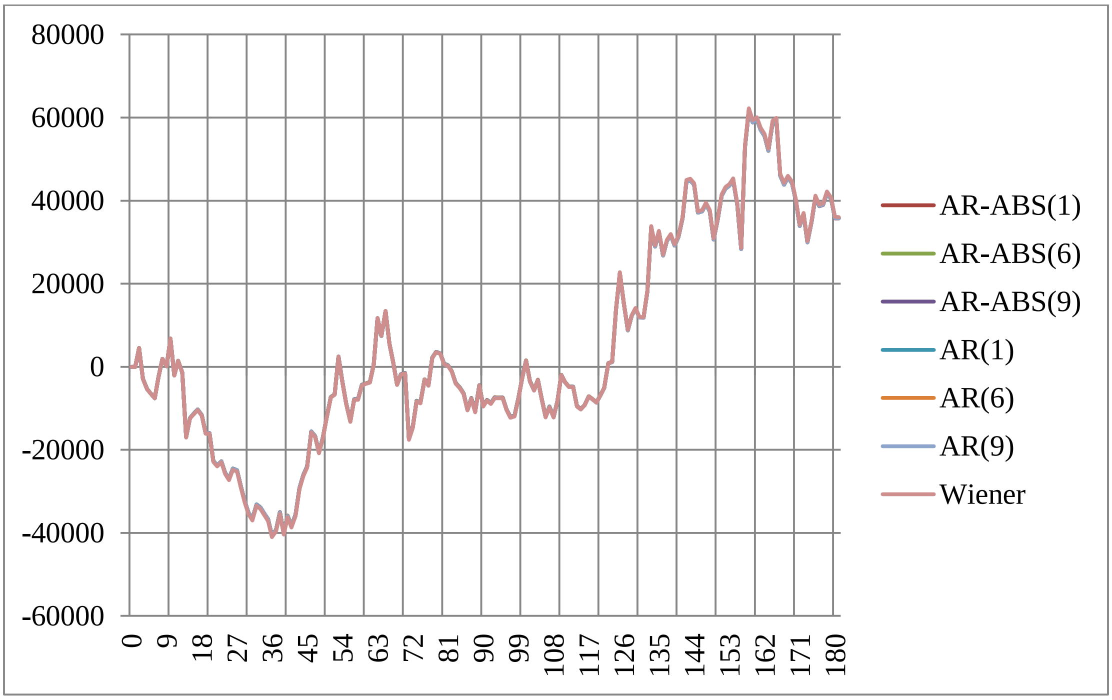

Fig. 41 *Daily profits of TR4*

Figure 42 shows average profits of TR6.

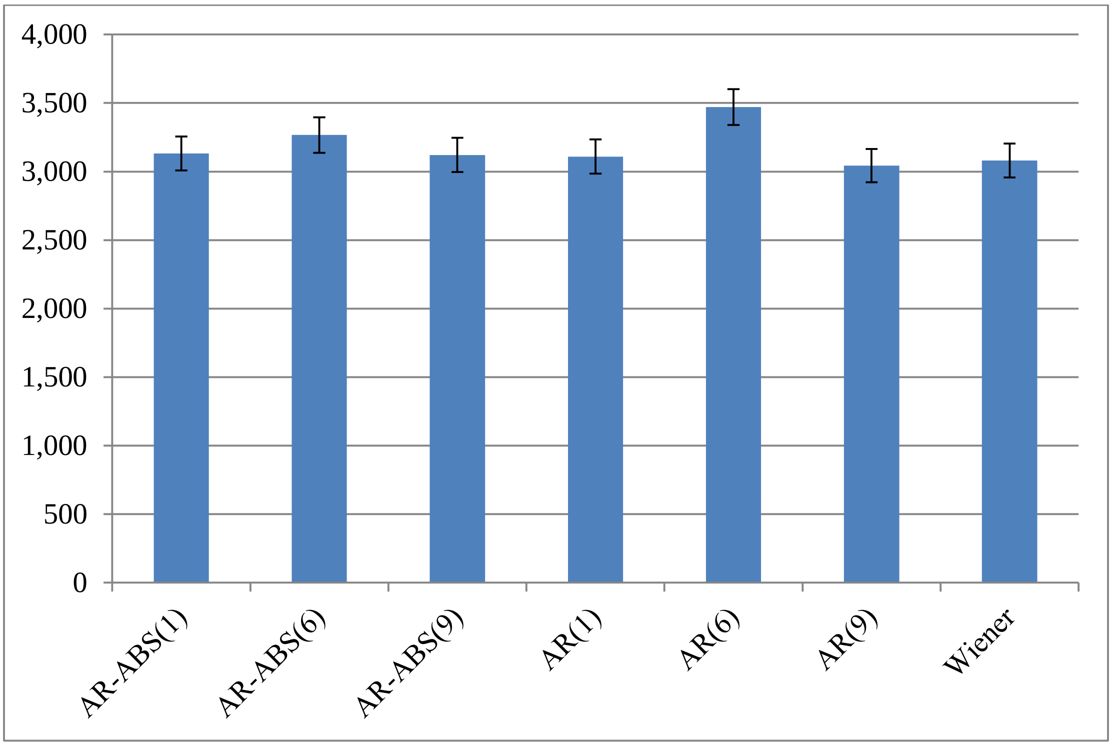

Fig. 42 *Average profits of TR5*

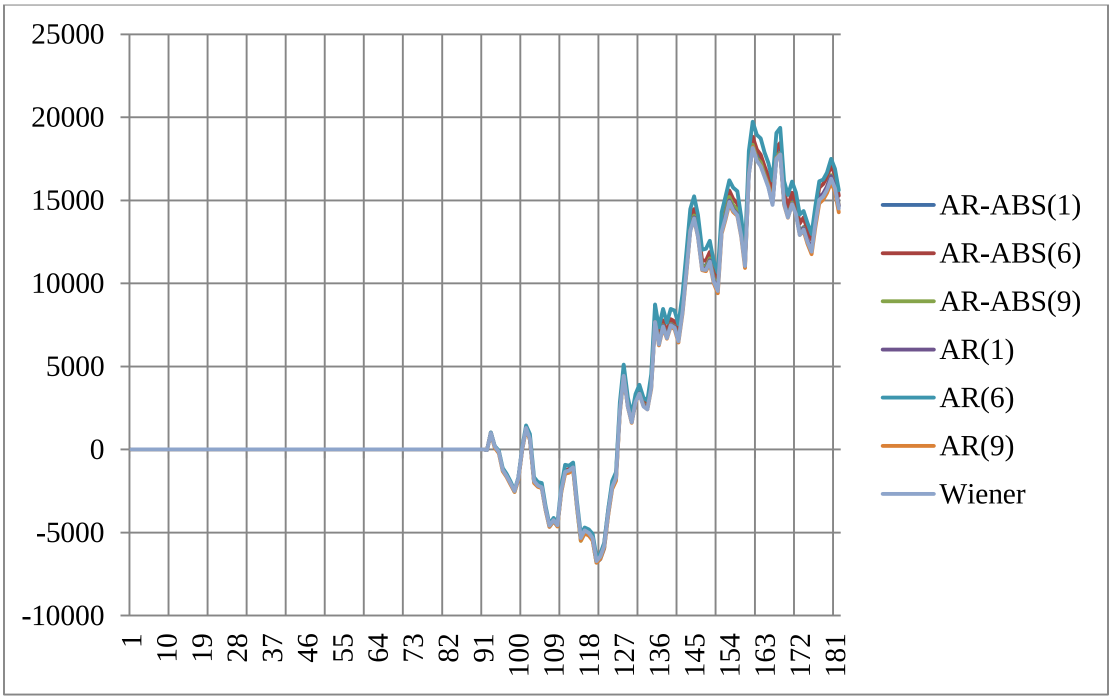

Fig. 43 *Daily profits of TR5*

Figure 43 shows daily profits of TR5. The horizontal line during the first 90 days represents the learning period when no trading was performed in accordance with the usual notion of long time strategies.

Figure 44 shows average profits of TR6.

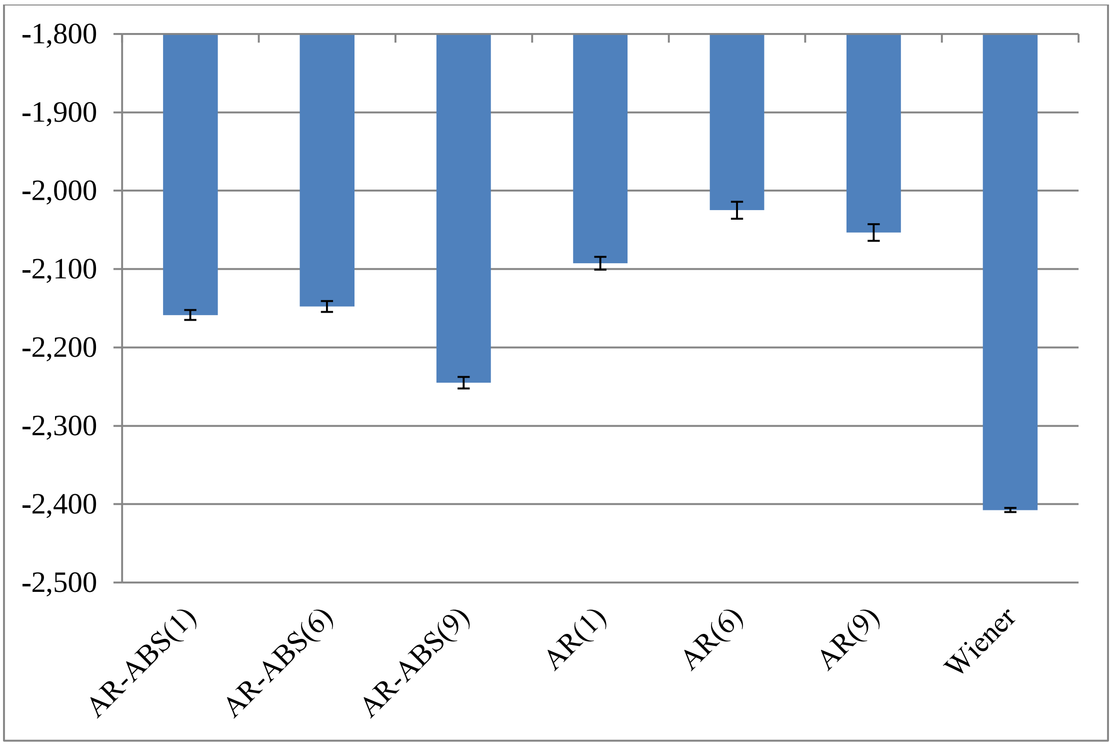

Fig. 44 *Average profits of TR6*

Figure 45 shows daily profits of TR6.

Unexpected result was that in contrast to other trading rules providing positive profits almost independently on prediction models, the application of TR6 shows losses which depends significantly on the prediction models. This means that both the patterns and values of profits strongly depend on the duration and time of data records.

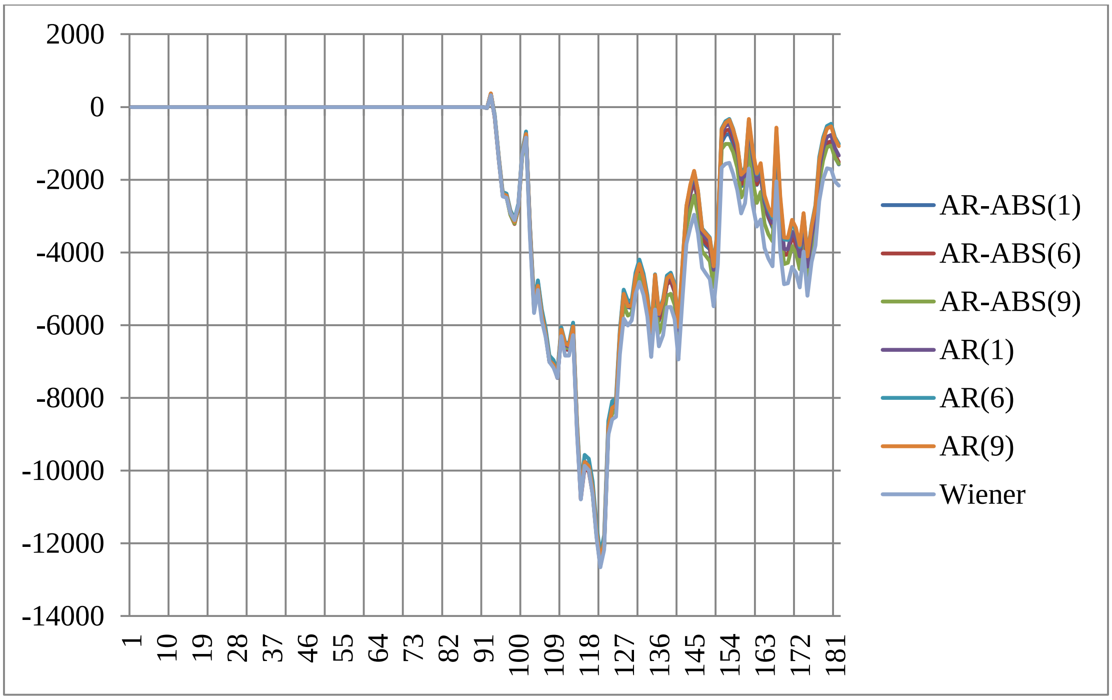

Fig. 45 *Daily profits of TR6*

4.7. Conclusions of Chapter 4

1. In most of the experiments, using all the autoregressive models, the minimal or close to minimal prediction error was achieved at parameter $p = 1$. Using the AR-ABS (p) models prediction errors were similar and close to the minimal for all parameters $p = 1,3,6,9$.
2. The experiments with both the historical and virtual financial data show that the minimal standard statistical prediction errors do not necessary provide maximal profits.
3. Both the statistical errors and average profits are very sensitive to data variations reflecting different economic conditions. However, the sensitivity of profits is greater.
4. The pattern of profits is different in different periods, representing different economic conditions while the patterns of prediction errors are similar.

5. In the post-crisis period, the correlation between the profits and prediction errors is positive in about half of the cases. In the virtual market, the positive correlation prevails. In both the post-crisis and virtual markets, the prices of most stocks grow.
6. The profitability of investments depends mainly on trading rules, so the optimization should be performed on the set of trading rules by the direct simulation of these rules using the corresponding stock-market models. This partly explains the weak correlation of profits and prediction accuracy.
7. Comparison of experimental results obtained using virtual and historical financial time series, shows that in non-stable post-crisis economical conditions, the historical results are similar to those of virtual ones.

Conclusion

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

1. In most of the experiments, using all the autoregressive models, the minimal or close to minimal prediction error was achieved at parameter $p = 1$.
2. Using the AR-ABS (p) models prediction errors were similar and close to the minimal for all parameters $p = 1,3,6,9$.
3. The experiments with both the historical and virtual financial data show that the minimal standard statistical prediction errors do not necessary provide maximal profits. Surprisingly, in the virtual markets, the positive correlation was observed. In the post-crisis recovery period, where the stock price graphs happened to be similar to the virtual ones, the positive correlation was in about half of experiments. In the stable economic conditions, the correlation was small but mainly negative, as expected.
4. Both the statistical errors and average profits are very sensitive to data variations reflecting different economic conditions. However, the sensitivity of profits is greater. The pattern of profits is different in different periods, representing different economic conditions while the patterns of prediction errors are similar.
5. The profitability of investments depends mainly on trading rules, so the optimization should be performed on the set of trading rules by the direct simulation of these rules using the corresponding stock-market models. This partly explains the weak correlation of profits and

prediction accuracy.

6. An important feature of the PORTFOLIO model is the multi-stock extension and a number of different trading rules which represent both the heuristics of potential investors and the well-known theoretical investment strategies This 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 improvemend comparing with traditional single-stock models with direct interaction of investment agents.
7. 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 (Krugman, 2000, 2008, 2009).
8. Comparison of experimental results obtained using virtual and historical financial time series shows that the results are similar in non-stable post-crisis economical conditions.
9. The PORTFOLIO model can be used as a tool to represent behavior of individual investor which 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 only available information is the historic data of real stocks.
10. There are many financial market models, but just a few stock exchange models. The well-known financial market models simulate interactions of independent agents trading a single stock. In contrast, the proposed model simulates the work of stock exchange trading many different stocks.
11. 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.

References

Arthanari, T. S., Dodge, Dr. Y. (1993). Mathematical Programing in Statistics. Wiley, New York. ISBN 978-0471592129.

Cochrane, J. H. (2005). Time Series for Macroeconomics and Finance. University of Chicago, Graduate School of Business, Chicago.

Darley V., Outkin A. (2004). A Nasdaq Market Simulation: Draft in Progress.

Fama, E. F. (1995). Random Walks in Stock Market Price. Financial Analysts Journal. Vol. 51, No. 1. p. 75–80.

Fishburn, P. C. (1964). Decision and value theory. Wiley, New York. ISBN 978-0471260653.

Fong, S., Si, Y.W., Tai, J. (2012). Trend following algorithms in automated derivatives market trading. Expert Systems with Applications. Vol. 39, Iss. 13. p. 11378-11390.

jFuzzyLogic (2012). An open source fuzzy logic library and FCL language implementation. Accessed at: <http://jfuzzylogic.sourceforge.net/html/index.html>

Krugman, P. (2000). Currency Crises (National Bureau of Economic Research Conference Report). The University of Chicago Press, Chicago and London. ISBN 978-0226454627.

Krugman, P. (2008). The Return of Depression Economics and the Crisis of 2008. W. W. Norton & Co, New York. ISBN: 978-0393337808.

Krugman, P. (2009). The increasing returns revolution in trade and geography. The American Economic Review. Vol. 99, No. 3. p. 561–571.

Landauskas, M., Valakevicius, E. (2011). Modelling of stock prices by the Markov chain Monte Carlo method. Intellectual economics. Vol. 5, No. 2(10). p. 244–256. ISSN 1822-8011.

MarketWatch. (2012). A free stock market game from MarketWatch. Accessed at: <https://www.marketwatch.com/game/>

Markowitz, H. M. (1952). Portfolio Selection. The Journal of Finance. Vol. 7, No. 1. p. 77-91.

Markowitz, H. M. (1959). Portfolio Selection: Efficient Diversification of Investments. Wiley, New York. ISBN 978-0300013726.

Marling, H., Emanuelsson, S. (2012). The Markowitz Portfolio Theory.

Merton, R. C. (1972). An analytic derivation of the efficient portfolio frontier. The Journal of Financial and Quantitative Analysis. Vol. 7, No. 4. p. 1851–1872.

Mockus, J. (2002). Stock exchange game model as an example for graduate level distance studies. Computer Applications in Engineering Education. Vol. 10, Iss. 4. ISSN 1061-3773. p. 229-237.

Mockus, J. (2010). On simulation of optimal strategies and Nash equilibrium in the financial market context. Journal of global optimization. Vol. 48, Iss. 1. p. 129-143.

Mockus, J. (2012). On simulation of the Nash equilibrium in the stock exchange contest. Informatica / Institute of Mathematics and Informatics. Vol. 23, No. 1. ISSN 0868-4952. p. 77-104.

Mockus, J., Eddy, W., Mockus, A., Mockus, L., Reklaitis, G. (1997). Bayesian Heuristic Approach to Discrete and Global Optimization: Algotithms, Visualization, Software, and Applications. Kluwer Academic Publishers, Boston/London/Dordrecht. ISBN 978-0792343271.

Mockus, J., Raudys, A. (1999). Comparison of ARMA and Multilayer Perceptron Based Methods for Economic Time Series Forecasting. Informatica / Institute of Mathematics and Informatics. Vol. 10, No. 2. ISSN 0868-4952. p. 231-244.

Mockus, J., Raudys, A. (2010). On the efficient-market hypothesis and stock exchange game model. Expert systems with applications. Vol. 37, Iss. 8. p. 5673–5681.

Ramanauskas, T., Rutkauskas, A. V. (2009). Building an artificial stock market populated by reinforcement-learning agents. Lietuvos Bankas, Working Paper Series. No. 6.

Raudys, S., Raudys, A. (2011). High frequency trading portfolio optimisation: Integration of financial and human factors. ISDA 2011: Intelligent systems design and applications: 11th international conference, November 22-24, 2011 - Córdoba, Spain. ISBN 978-1457716751. p. 696-701.

Raudys, S., Raudys, A. (2012). Three decision making levels in portfolio management. CIFEr 2012: IEEE conference on computational intelligence for financial engineering & economics, March 2012, New York, USA: proceedings. ISBN 978-1467318037. p. 220-227.

Rolf, S., Sprave, J., Urfer, W. (1997). Model Identification and Parameter Estimation of ARMA Models by Means of Evolutionary Algorithms. In Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering. IEEE Press. p. 237-243.

Sharpe, W. F. (1966). Mutual fund performance. The Journal of Business. Vol. 39, No. 1. p. 119–138.

Sharpe, W. F. (1994). The Sharpe Ratio. The Journal of Portfolio Management. Vol. 21, No. 1. p. 49–58.

Shen, S., Jiang, H., Zhang, T. (2012). Stock Market Forecasting Using Machine Learning Algorithms. Stanford University.

StockTrak. (2012). Stock Market Simulation Software to Practice Trading. Accessed at: <http://www.stocktrak.com/public/products/Stock-Trak.aspx>

Voss, M. S., Feng, X. (2002). Arma Model Selection Using Particle Swarm Optimization and AIC Criteria. In Proceedins of 2002 IFAC Conference, 15th Triennial World Congress, Barcelona, Spain.

Wilmott, P. (2007). Introduces Quantitative Finance. Second Edition. Wiley, New York. ISBN 978-0470319581.

Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex system anddecision process. IEEE Transactions on Systems, Man, and Cybernetics. Vol. SMC-3, No. 1. p. 28–44.

Zuo, Y., Kita, E. (2012). Stock price forecast using Bayesian network. Expert Systems with Applications. Vol. 39, iss. 8. p. 6729-6737.

List of Publications

Articles in the reviewed scientific periodical publications

A 1. 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.

A 2. 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.

A 3. 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.

A 4. 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.

Appendices

Appendix A. Information on Independent Application, Testing and Verification of the PORTFOLIO Model

The guide is for Windows environment. The procedure is similar in Mac and Linux.

The Database

Step 1. Install XAMPP. To do this, download the free XAMPP software, and, after some “Next” steps, select Apache, MySQL, PHP, phpMyAdmin components:

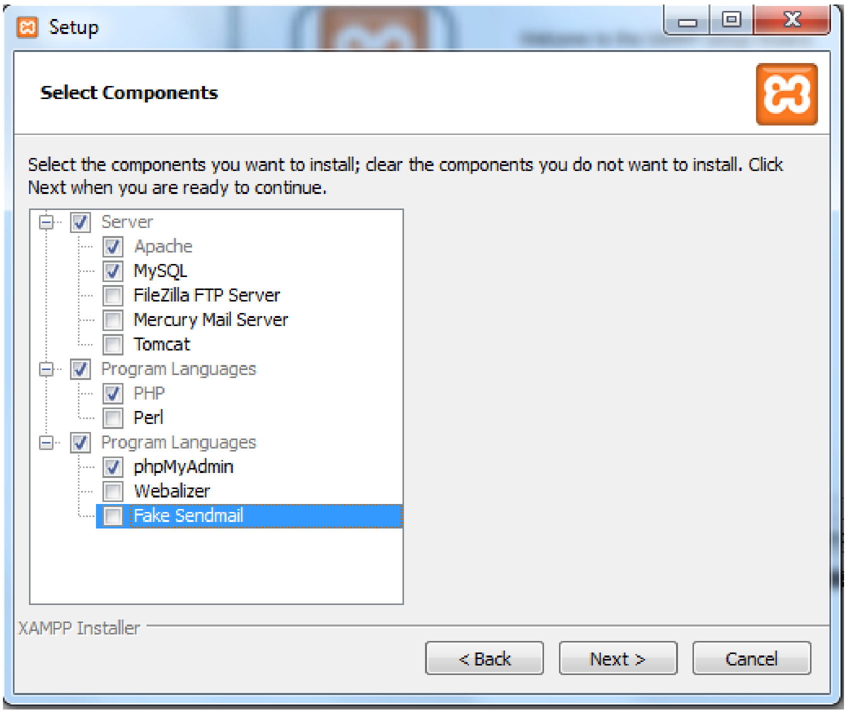

Fig. 1 *Selection of XAMPP components*

After some additional "Next" steps check "Finish" to open the XAMPP Control Panel:

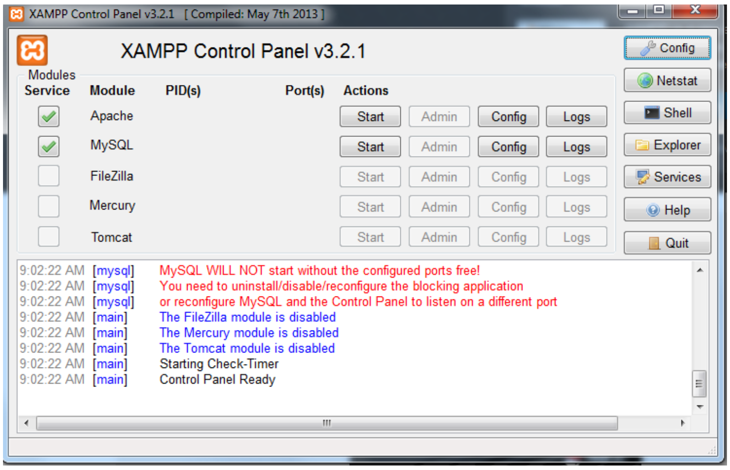

Fig. 2 *Starting Apache and MySQL*

By checking "Netstat" button, provide that ports 80 and 3306 would be free. Otherwise, change the ports in the Apache and MySQL settings by checking the corresponding "Config" button and editing "httpd.config" file:

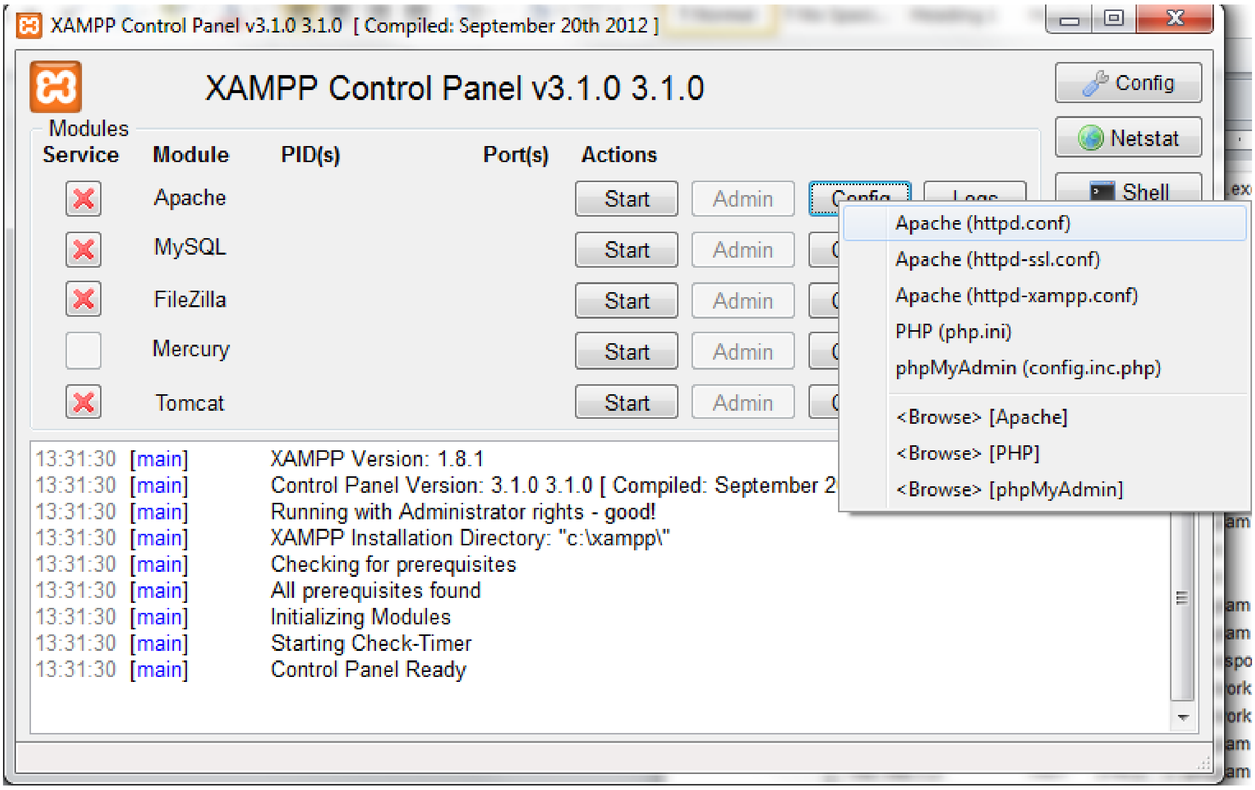

Fig. 3 *Configuring Apache and MySQL ports*

If port 80 is not free, then in the line “Listen 80” change the port number to the free one, for example 90.

The SSL port can be changed by editing the line “Listen 443” in the “httpd-ssl.conf” file.

MySQL port can be changed by editing the line “Port = 3306”.

Step 2. Start phpMyAdmin. If the ports were not changed start http://localhost/phpmyadmin/. If, for example, the port 80 was changed to 90, then start http://localhost:90/phpmyadmin/. This operation opens the window in Figure 4.

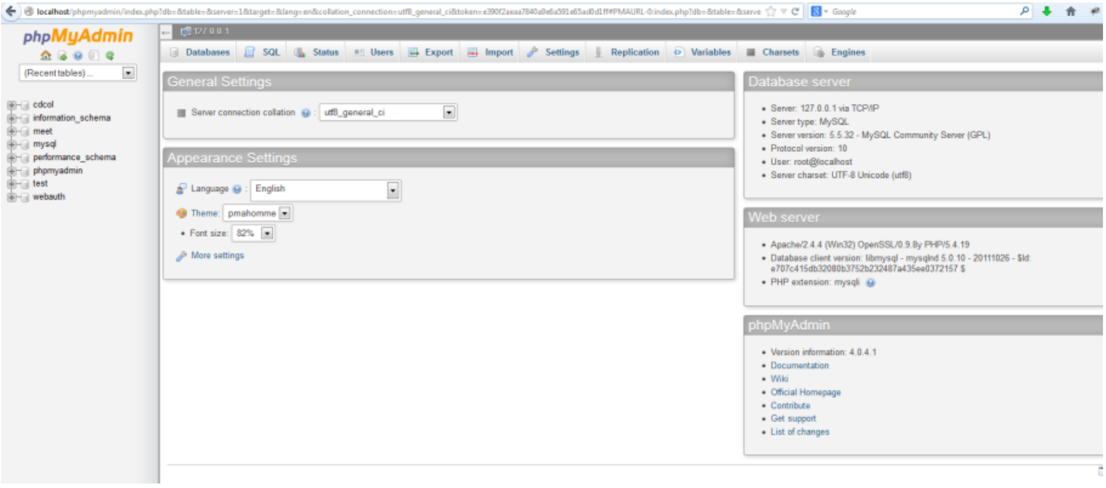

Fig. 4 *Opening phpMyAdmin window*

Step 3. Create the data base, for example "experiment". Check "Database", write the name "experiment", and select corresponding code:

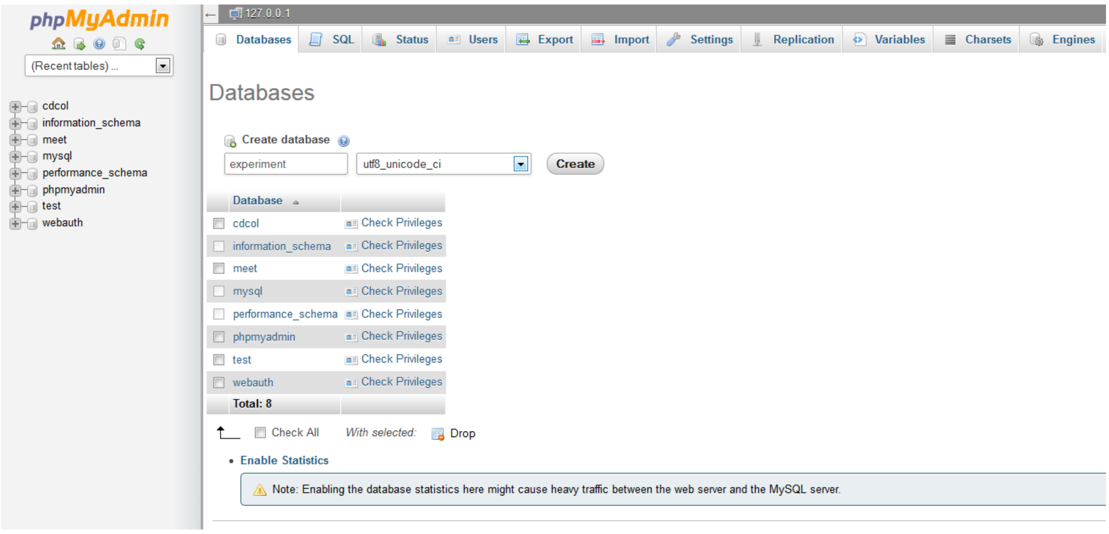

Fig. 5 *Creating the database "experiment"*

Check "Create", and see the new data base in the left side. If no data base is seen, check for the errors and repeat the process.

Step 4. Download the Java archive "stock.zip" using web-sites http://getweb.lt/igor/stock.zip, or http://optimum2.mii.lt/~jonas2, or http://fmf.vgtu.lt/~mockus, or http://mockus.org/optimum. In the last three sites the archive "stock.zip" is in the section Global Optimization, in the task

PORTFOLIO.

Extract the “stock.zip” archive and open the applet “index.html” by a browser with full Java support. The Java support may be provided by enabling the browsers Java plugin and by setting Java security policy in the following way:

Open Java Control Panel and set Java security level as shown in the next Figure. A way to open this panel in Windows 8.1 is by running the command C:\program files\java\jdk1.7.0_51\jre\bin\javacpl.exe as an administrator.

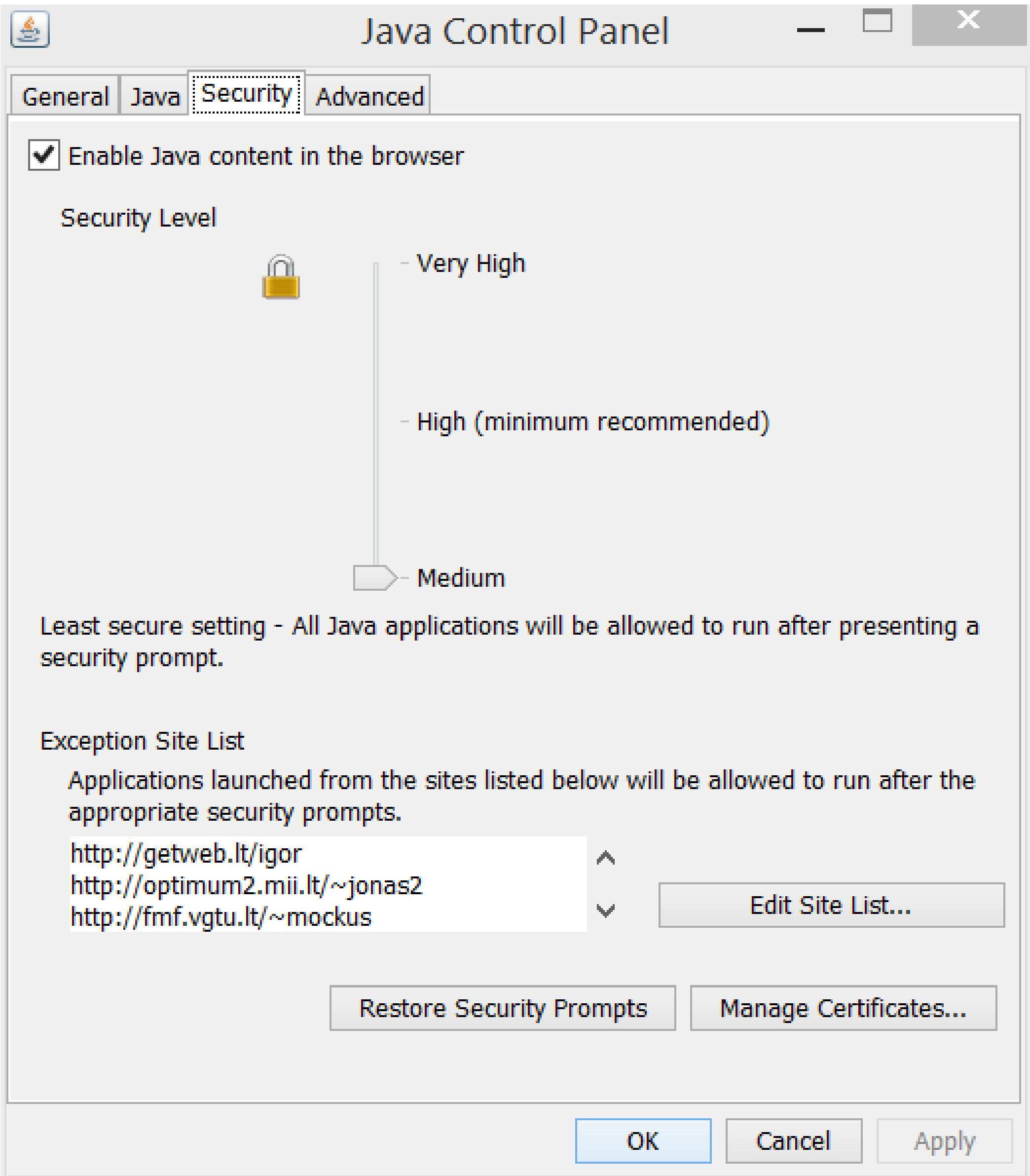

Fig. 6 *Configuring Java security*

In Figure 6 the security level is set to “medium” and the three sites are added for special permissions. Java is developing, so some new security setting may be needed.

Step 5. Select the number of stocks and other initial data. In Figure 7, four stocks are selected, all four of virtual market (generated data).

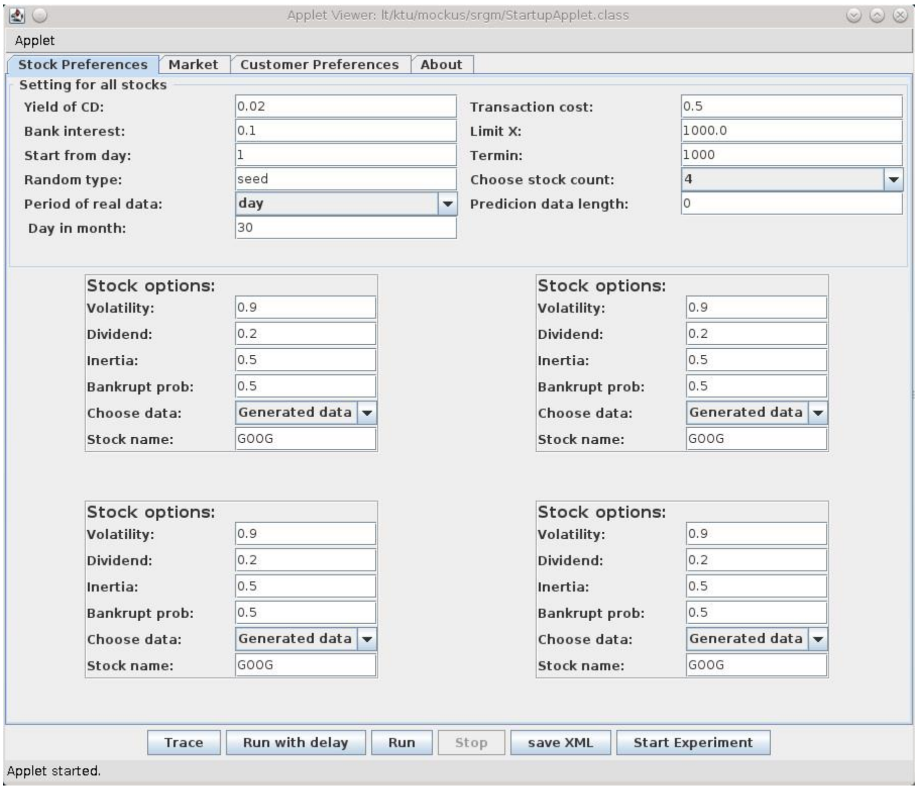

Fig. 7 *Setting the number and parameters of stocks*

Select the investors and trading strategies. In Figure 8 two investors were selected.

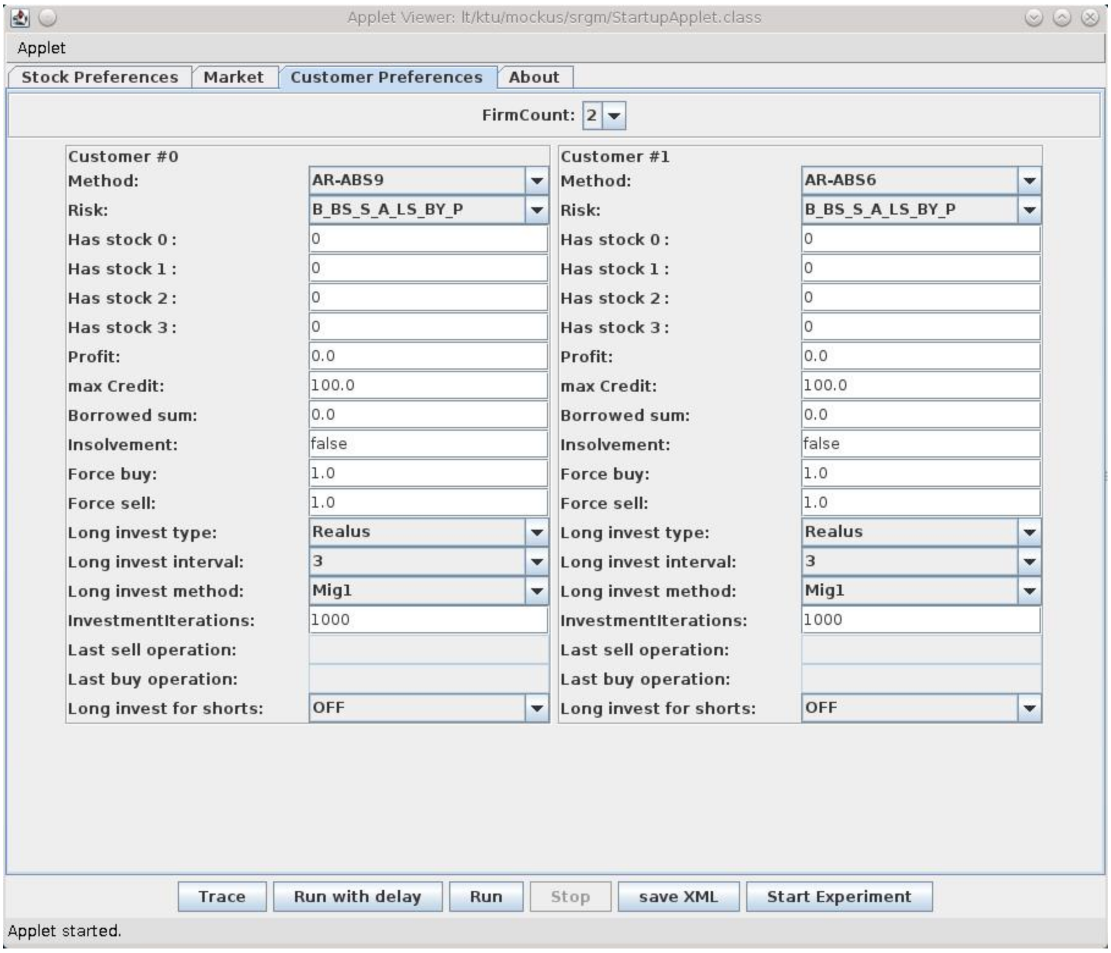

Fig. 8 *Setting the number of investors and their and parameters*

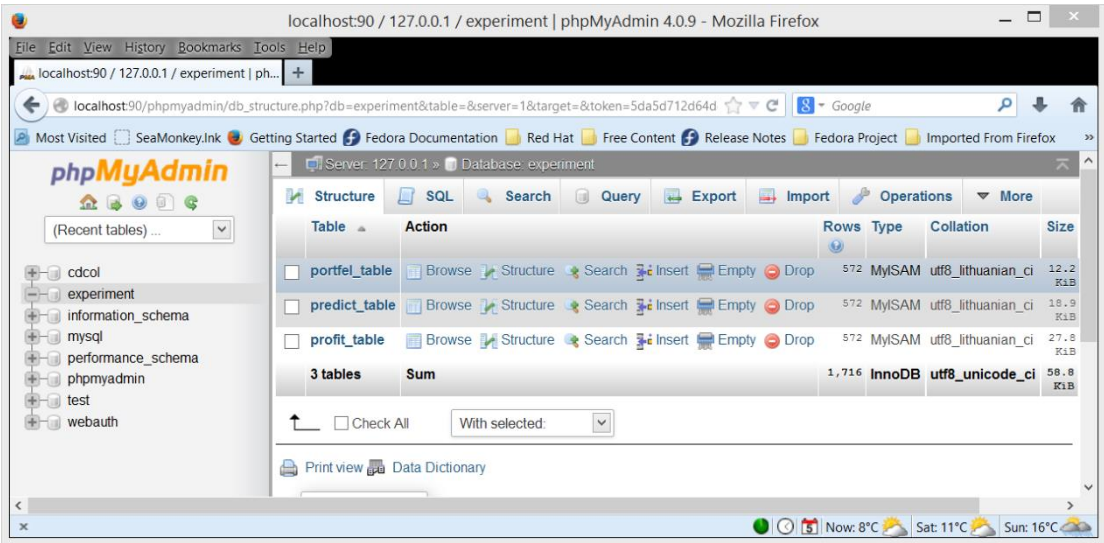

Fig. 9 *The experiment window*

Step 6. Start experiment by checking the corresponding button in the lower-right of Figure 9. If no mistakes, the Figure 8 will appear.

After the experiment is finished, the data can be extracted by standard means of SQL.

The Java Code

Java code is in the extracted archive “stock.zip” in the folder “Source Packages”. The Figure 10 shows the Java class “StockTradeThread” in the source folder “lt.ktu.mockus.srgm” opened as the NetBeans project “stock”.

Additional trade rules can also be included. Recompile, if needed, by “clean and build” and start new applet by opening “index.html” using a browser with Java support.

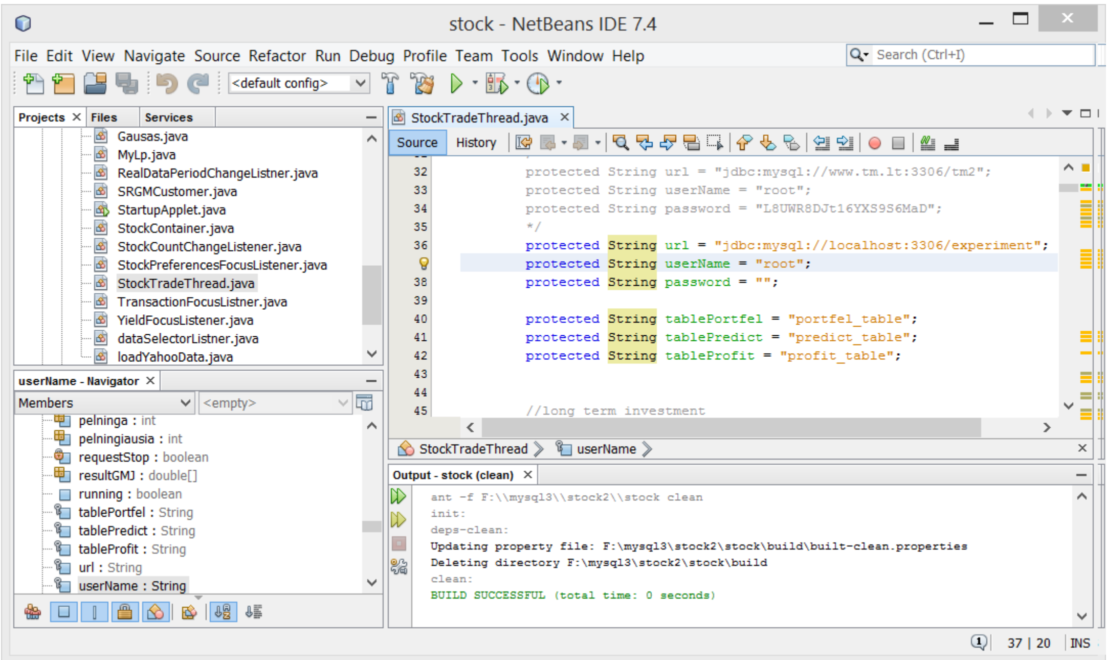

Fig. 10 *A fragment of Java code in the NetBeans project*

The software of the complete model is not easy for understanding. However, the means for the calculation of profits and prediction errors are very simple. The profit is calculated using the following Java class as a difference between selling and buying prices minus transaction costs and bank charges. The prediction errors, for example MAE, is the absolute value of difference between the predicted and observed stock price values. It is calculated by the MySQL query shown in the Example 1.

```
double vartotojoVisuAkcijuVerte = 0;
for( int ccc = 0; ccc < this.applet.MAX_STOCK_COUNT; ccc++ ) {
vartotojoVisuAkcijuVerte += ((Integer)customer.iN.get(ccc)).intValue()
* customer.currentPrice.get(ccc);
}
customer.dProfit = customer.C0 - customer.C0_ - customer.B + vartotojoVisuAkcijuVerte;
```

In the **Customer** class, a field **dProfit** defines profit at the current moment using current market prices multiplied by the number of stocks.

Here, **C0_** is initial funds, **C0** is invetors cash, **B** is borrowed money, and **vartotojoVisuAkcijuVerte** is the value of all stocks belonging to the investor.

Three examples of SQL query follows. Each example should be used separatelly.

Example 1 for prediction errors:

```
SELECT `t1`.`stock`, `t1`.`strategy`, (
SUM(ABS(t1.kaina - t1.prog)) / SUM(t1.kaina))MAE,
(SQRT(SUM((t1.kaina-t1.prog)*(t1.kaina-t1.prog)))
/ SUM(t1.kaina)) SE
FROM (
SELECT `day`, `stock`, `strategy`, AVG(`price`) kaina, AVG(`predict`) prog
FROM `predict_table`
GROUP BY `day`, `stock`, `strategy`) `t1`
GROUP BY `t1`.`strategy`,`t1`.`stock`
```

Example 2 for portfolio:

```
SELECT `strategy`, `stock`, AVG(`stockCount`)
FROM `portfel_table`
GROUP BY `strategy`, `stock`
LIMIT 0, 30
```

Example 3 for profits:

```
SELECT `strategy_predict`,
DAY, AVG(profit)
FROM `profit_table`
GROUP BY `strategy_predict`,
DAY
LIMIT 0, 30
```

Igor Katin

ON DEVELOPMENT AND INVESTIGATION
OF STOCK-EXCHANGE MODEL

Doctoral Dissertation

Technological Sciences,
Informatics Engineering (07T)

Igor Katin

AKCIJŲ BIRŽOS MODELIO
SUDARYMAS IR TYRIMAS

Daktaro disertacija

Technologijos mokslai,
Informatikos inžinerija (07T)