ISSN 1822-8011 (print)
ISSN 1822-8038 (online)
INTELEKTINĖ EKONOMIKA
INTELLECTUAL ECONOMICS
2009, No. 2(6), p. 83–95

AGENT-BASED STOCK MARKET SIMULATION MODEL

Aleksandras Vytautas RUTKAUSKAS

Vilnius Gediminas Technical University
Saulėtekio ave. 11, LT-10223 Vilnius, Lithuania
E-mail: ar@vv.vgtu.lt

Tomas RAMANAUSKAS

Vilnius Gediminas Technical University
Bank of Lithuania
Totorių str. 4, LT-01121 Vilnius, Lithuania
E-mail: tramanauskas@lb.lt

Abstract. In this paper we propose an artificial stock market model based on the interaction of heterogeneous agents whose forward-looking behaviour is driven by the reinforcement learning algorithm combined with an evolutionary selection mechanism. We use the model for the analysis of market self-regulation abilities, market efficiency and determinants of emergent properties of the financial market. Novel features of the model include a strong emphasis on the economic content of individual decision-making, the application of the Q-learning algorithm for driving individual behaviour, and rich market setup. A parallel version of the model which is based on the research of current changes in the market as well as on the search for newly emerged consistent patterns and which has been repeatedly used for optimal decisions' search experiments in various capital markets is presented.

JEL classification: G10, G11, G14.
Keywords: agent-based financial modelling, artificial stock market, complex dynamical system, emergent properties, reinforcement learning.
Reikšminiai žodžiai: agentais pagrįstas finansinis modeliavimas, dirbtinė akcijų rinka, kompleksinė dinaminė sistema, kylančios savybės, skatinamasis mokymasis.

Introduction

In this paper we develop an artificial stock market (ASM) model, which could be used to examine some emergent features of a complex system comprised of a large number of heterogeneous learning agents that interact in a detail-rich and realistically designed environment. This version of the model is not calibrated to empirical data; therefore, at this stage the main aim of the present research is to offer, implement and test some new ideas that could lay ground for a robust framework for the analysis of financial market processes and their determinants. We believe that the model does offer an interesting framework for the structured analysis of market processes without abstracting from relevant and important features, such as an explicit trading process, regular dividend payouts, trading costs, agent heterogeneity, dissemination of experience, competitive behaviour, agent prevalence and forced exit, etc. Of course, some of these aspects have already been incorporated in the existing agent-based financial models. However, the lack of a widely accepted fundament in this area of modelling necessitates the individual and largely independent approach, which is pursued in this study. What is more, an alternative, from this viewpoint, decisions management system in capital markets is analysed; the system is based on certain assumptions about the continuity of capital market behaviour and on newly formed features and can be efficiently used in various capital markets during the global financial crisis.

One of the most interesting features of ASM modelling is a relatively detailed modelling of the decision processes. In our view, agent-based models developed to deepen our understanding of the real world financial processes can only be fully utilised if a strong emphasis is put on the economic content of the model, i.e. individual behaviour and market structure must be based on clear and economically sound principles. Importantly, agents in our model exhibit economically appealing and forward-looking behaviour, which is based on adaptive learning or, more specifically, a combination of reinforcement learning (Watkins' Q-learning) and evolutionary change. To our knowledge, this is one of the first attempts to incorporate Q-learning algorithm into ASM models.

By conducting simulation experiments in this model, we aim to address some specific questions, such as market self-regulation abilities, the congruence between the market price of the stock and its fundamentals (the market efficiency issue), the importance of intelligent individual behaviour and interaction for market efficiency and functioning, and the relationship between stock prices and market liquidity.

1. Related Literature Review

The ASM research area is relatively new but there is a growing body of literature on the subject. There is a clear lack of a comprehensive literature review and classification of the existing models. Some popular models and ASM modelling principles are discussed in LeBaron (2006), while Samanidou et al. (2007) present a review of some agent-based financial models with the emphasis on econophysics. At the heart of ASM models is the interaction of heterogeneous agents which leads to complex systemic behaviour and emergent systemic properties. There are two broad classes of ASM models, namely, models based on agents' hard-wired behavioural rules (see Kim and Markowitz (1989), Sethi and Franke (1995), Lux (1995)) and models supporting systemic adaptation. The most prominent example of the latter category is the Santa Fe ASM model developed by Arthur et al. (1997) (also see Beltrati and Margarita (1992), Lettau (1997), LeBaron (2000), Tay and Linn (2001)). See Ramanauskas (2009) for a general discussion of agent-based financial modelling and the abovementioned models. In many models systemic adaptation is usually warranted by evolutionary algorithms, whereas individual agents' behaviour is very stylised and based on economic consideration directly. In contrast, in modelling financial market processes, we put a strong emphasis on individual behaviour and economic reasoning.

Lithuanian researchers have been interested in stock market analysis and modelling of investment strategies for more than a decade but the specific area of ASM modelling has not been systematically researched and, to our knowledge, no full-fledged artificial stock market models have been developed by Lithuanian researchers. Studies of investment strategies are conceptually most closely related to our research. Some of the most important studies of investment strategies must be briefly mentioned. Plikynas et al. (2002) made early attempts to use neural networks in stock market forecasting. Nenortaitė and Simutis (2005), Nenortaitė (2007) apply artificial neural networks and particle swarm optimization algorithms to develop stock trading strategies based on historical stock performance. Simutis and Masteika (2004) use fuzzy neural networks and evolutionary programming methods for creating expert systems for stock trading. Rutkauskas and Stasytytė (2008) implement risk stratification procedure to augment the standard risk-return paradigm of investment risk taking. Stankevičius (2001) uses the idea of self-organising maps for the formation of investment portfolios. A notable contribution by Plikynas (2008) is the development of a multi-agent trading system based on competing heterogeneous neural network strategies.

2. Description of the ASM Model

The present ASM model does not fully abstract from many important features of real financial markets that are usually excluded both from standard financial models and other ASMs. For example, just like in the real world financial markets, agents in this ASM model do not know the "true model" but try instead to adapt in the highly uncertain environment; they exhibit bounded rationality, non-myopic forward-looking behaviour, as well as diversity in experience and skill levels; the trading process is quite realistic and detailed; dividends are paid out in discrete time intervals and the importance of dividends as a fundamental force driving stock prices is explicitly recognised. The proposed ASM model embodies some new ideas about financial market modelling and provides an interesting generative explanation of the prolonged periods of over- and under-valuation. In this section we present the architecture of the artificial stock market in detail.

2.1. General Market Setting and Model's Main Building Blocks

The artificial stock market is populated by a large number of heterogeneous reinforcement-learning investors. Investors differ in their financial holdings, expectations regarding dividend prospects or fundamental stock value. This ensures diverse investor behaviour even though the basic principles governing experience accumulation are the same across the population. The very basic description of agents' behavioural principles can be as follows. All agents forecast an exogenously given, unknown dividend process and base their estimates of the fundamental stock value on dividend prospects. These estimates are intelligently adjusted to attain immediate reservation prices. Agents explore the environment and accumulate the experience with the aim of maximising long-term returns on their investment portfolios but there are no optimality guaranties in the context of high uncertainty and complex interaction of agents.

As usual in financial market modelling, the modelled financial market is very simple. Only one, dividend-paying stock (stock index) is traded on the market. Dividends are generated by an exogenous stochastic process unknown to the agents, and they are paid out in regular intervals. The number of trading rounds between dividend payouts can be set arbitrarily, which enables interpretation of a trading round as a day, a week, a month, etc. Paid out dividends and funds needed for liquidity purposes are held in private bank accounts and earn constant interest rates, whereas liquidity exceeding some arbitrary threshold is simply removed from the system (e.g. consumed). Borrowing is not allowed. Initially, agents are endowed with arbitrary stock and cash holdings, and subsequently in every trading round each of them may submit a limit order to buy or sell *one unit* of stock, provided, of course, that financial constraints are non-binding. Trading takes place via the centralised exchange.

To facilitate the detailed model exposition, it is useful to break the model into a set of economically meaningful processes, though some of them are inter-related in complex ways. We will discuss these logical building blocks in the following subsections.

2.2. Forecasting Dividends

Expected company earnings and dividend payouts are the main fundamental determinants of the intrinsic stock value. We assume that all agents make their private forecasts of dividend dynamics. Dividend flows are generated by an unknown, potentially non-stationary data generating process specified by a modeller. The only information, upon which agents can base their forecasts, is past realisation of dividends, and agents know nothing about the stationarity of the data generating process. Hence, they are assumed to form adaptive expectations, augmented with the reinforcement learning calibration. We also allow for the possibility to improve a given agent's forecasting ability by probabilistic imitation of more successful individuals' behaviour.

Agents start with finding some basic reference points for their dividend forecasts. The exponentially weighted moving average (EWMA) of realised dividend payouts can be calculated as follows:

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

Here d_y denotes dividends paid out in period y (year) and λ_1 is the arbitrary smoothing factor (the same for all agents), which is a real number between 0 and 1. The subscript i on the averaged dividends in equation (1) indicates that they vary across individual agents. The differences arise due to different arbitrarily chosen initial values but over time, however, these exponential averages converge to each other. Also note that dividend payouts can be arbitrarily less frequent than stock trading rounds, e.g. if one trading period equals one month, dividends may be scheduled to be paid out every twelve periods and in equation (1) one time unit would be one year.

Exponential moving averages would clearly be unacceptable estimates of future dividends in a general case. Hence, their function in this model is twofold. Firstly, they provide a basis for further "intelligent" refinement of dividend forecasts, i.e. these moving averages are multiplied by some adjustment factors calibrated in the process of the reinforcement learning. And secondly, forecasting dividends relative to their moving averages, as opposed to forecasting dividend levels directly, make forecasting environment more stationary, which facilitates the reinforcement learning task.

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

$$E(d_{i,y+n}) = d_{i,y}^{EWMA} \cdot a_{i,y}^{div}, \qquad (2)$$

where $a_{i,y}$ is the 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. Individual forecasts for periods $y+1, \ldots, y+n$ formed in periods $y-n+1, \ldots, y$, respectively, are stored in the program and used for determining individual estimates of the fundamental stock value.

2.3 Estimating Fundamental Stock Value and Reservation Prices

Quite similarly to the dividend forecasting procedure, agents' estimation of the intrinsic stock value is a two-stage process. It embraces the formation of initial estimates of the fundamental value, based on discounted dividend flows, and ensuing intelligent adjustment grounded on agents' interaction with environment. We refer to this refined fundamental value as the reservation price.

The initial evaluation of the future dividend flows is a simple discounting exercise. To calculate the present value of expected dividend stream, the constant interest rate is used as the discount factor. Moreover, beyond the forecast horizon dividends are assumed to remain constant. Under these assumptions, individual estimates of the present value of expected dividend flows are

$$v_{i,y}^{fund} = d_y + E\left(\frac{d_{i,y+1}}{1+\bar{r}} + \ldots + \frac{d_{i,y+n}}{(1+\bar{r})^n} + \frac{d_{i,y+n}/\bar{r}}{(1+\bar{r})^{n+1}}\right), \quad (3)$$

where $\bar{r}$ is the constant interest rate. The last term in this equation is simply the discounted value of the infinite sum of steady financial inflows. These present value estimates are subject to further refinement.

To avoid excessive volatility of the estimates of the discounted value of dividend stream, they are again smoothed by calculating the exponentially weighted moving averages:

$$v_{i,y}^{EWMA} = \lambda_2 \cdot v_{i,y}^{fund} + (1-\lambda_2)v_{i,y-1}^{EWMA}. \quad (4)$$

The role of these averages is very similar to that of the averaged dividends in the dividend forecasting process, namely, to provide some background for the reinforcement learning procedure and (partially) stationarise the environment in which agents try to adapt.

The second stage in the estimation of the individual reservation prices of the stock is the calibration based on the reinforcement learning procedure. For this we have to switch to the different time frame (in the base version of the model it is assumed that dividends are paid out annually, whereas agents can trade once per month). In a given trading round t, individual reservation prices $v_{i,t}^{reserve}$ are obtained from equation (4) by multiplying exponentially smoothed estimates of fundamental value by individual price adjustment factors, $a_{i,t}^{p}$:

$$v_{i,t}^{reserve} = v_{i,t}^{EWMA} \cdot a_{i,t}^{p}. \quad (5)$$

In this context the individual reservation price is understood as an agent's subjective assessment of the stock's intrinsic value that prompts immediate agent's response (to buy or sell the security).

2.4. Making Individual Trading Decisions

Having formed their individual beliefs about the fundamental value of the stock price, agents have to make specific portfolio rebalancing decisions. In principle, they weight their own assessment of the stock against market perceptions and make orders to buy (sell) one unit of the underpriced (overpriced) stock at the price that is expected to maximise their wealth at the end of the trading period. We give a more detailed description of these processes below.

The individual reservation price reflects what investors think the stock price should be worth. If the last period's average market price p_{t-1} is less than agent i's reservation price today, the agent is willing to buy stock and pay at most $v_{i,t}^{reserve}$. Conversely, if the prevailing market price is higher than the agent's perceived fundamental, the agent is willing to sell it at $v_{i,t}^{reserve}$ or higher price. So its decision rule is like this:

If $v_{i,t}^{reserve} > p_{t-1}$ *and* $m_{i,t}^{0}$ *is sufficient* $\rightarrow$ *submit limit order to buy 1 share at price* $p_{i,t}^{q}$

if $v_{i,t}^{reserve} < p_{t-1}$ *and* $h_{i,t}^{0} > 0$ $\rightarrow$ *submit limit order to sell 1 share at price* $p_{i,t}^{q}$

otherwise, make no order.

Here $h_{i,t}^{0}$ and $m_{i,t}^{0}$ denote, respectively, agent i's stock holdings (i.e. number of owned shares) and cash balance at the beginning of a trading round, $p_{i,t}^{q}$ is the quoted price to be determined below.

Agents, of course, aim at getting most favourable prices for their trades but they must take into account the fact that better bid or ask prices are generally associated with smaller probabilities of successful trades. The assumption that each agent is allowed to trade only one unit of stock in a given trading round has a very useful implication in this context – the probabilities of successful trades at all possible prices faced by a buyer and a seller can be loosely interpreted as the supply and demand schedules, respectively. So we further assume that these supply and demand schedules are estimated by the exchange institution from past trading data and constitute public knowledge.

Estimated probabilities of successful trades at given (relative) price quotes are computed as follows. Simply put, these estimated probabilities should indicate chances of successful trading at prices that are "high" or "low" relative to the prevailing market price (i.e. last period's average price). So the probability of a successful trade for a given price quote (relative to the benchmark price) is calculated from the past trading rounds as a fraction of successfully filled buy (sell) orders out of all submitted orders to buy (sell) at that price. Unfortunately, due to computational constraints the number of agents and successful trades is not sufficiently high to obtain reliable estimated probabilities in this straightforward way. For this reason we employ the following three-step procedure:

i) estimates of probabilities of successful buy and sell orders for every price quote are smoothed over time by computing exponential moving averages;

ii) if there are no orders to buy or sell at a given price at time t, the exponential moving average estimates of successful trade probabilities are left unchanged from the t–1 period;

iii) the scattered estimates are fitted to a simple cross-sectional regression line (with its values restricted to lie in the interval between 0 and 1) to ensure that the sets of successful trade probabilities retain meaningful economic properties.

As a result, we get a nice upward-sloping line, which represents probabilities of successful buy orders for each possible price quote, and a downward-sloping line for the sell orders case. At this stage agents have all the components needed to choose prices that give them highest expected wealth at the end of the trading round. Firstly, agent i estimates the expected end-of-period stock holdings (i.e. the number of shares) for each possible price quote j:

$$E(h^1_{i,j,t}) = h^0_{i,t} + E(q_{i,j,t}) \cdot b_i \text{ for all } j \quad . \qquad (6)$$

Here $E(q_{i,j,t})$ denotes expected number of shares to be bought or sold by agent i at any quotable price j (as was explained above, these numbers lie in the closed interval between 0 and 1). The indicator variable b_i takes value of 1 if the agent is willing to buy the stock or -1 if the agent is willing to sell the stock.

Similarly, agent i's expected end-of-period cash holdings for each possible price quote j are

$$E(m^1_{i,j,t}) = m^0_{i,t} + E(q_{i,j,t}) \cdot x_{j,t} \cdot (-b_i - c) + E(h^1_{i,j,t}) \cdot E(d_{i,t}) \text{ for all } j. \qquad (7)$$

Here $x_{j,t}$ denotes possible price quote j, c is the fractional trading cost, and $E(d_{i,t})$ denotes the expected dividends, which are to be paid out following the trading round (this term equals zero in between the dividend payout periods). It is important to note that the interest on spare cash funds is paid, as well as excess liquidity (cash holdings above some pre-specified amount needed for trading) is taken away at the beginning of the trading period. This is reflected in $m^0_{i,t}$. Dividends are paid out for those agents that hold stocks after the trading round, as can be seen from equation (7).

Finally, agent i's expected end-of-period stock holdings are valued at individual reservation price and each agent calculates its expected end-of-period wealth for every possible price quote:

$$E(w^1_{i,j,t}) = E(h^1_{i,j,t}) \cdot v^{reserve}_{i,t} + E(m^1_{i,j,t}). \qquad (8)$$

Hence, agent i's quoted price, p^q_i, is the price that is associated with the highest expected wealth at the end of the trading round:

$$p^q_i = \arg\max_{x_i} E(\vec{w}^1_{i,t}). \qquad (9)$$

If several price quotes result in the same expected wealth, the agent chooses randomly among them. It is also important to note that in the process of the reinforcement learning, agents are occasionally forced to take exploratory actions. In those cases exploring agents choose prices from the quote grid in a random manner.

Market price determination and actual trading take place through a centralised stock exchange. The trading mechanism basically is the double auction system, in which both buyers and sellers contemporaneously submit their competitive orders to implement their trades. Agents are assumed to have no knowledge of individual market participants' submitted orders. The centralised stock exchange also produces a number of trading statistics, both for analytical and computational purposes. These statistics include the market price, trading volumes and volatility measures. The market price in a given trading period is calculated as the average traded price.

2.5. Learning and Systemic Adaptation in the Model

We assume that the agents' behaviour is driven by reinforcement learning since these learning algorithms borrowed from the literature on machine learning seem to be conceptually suitable for modelling investor behaviour (see, Bertsekas and Tsitsiklis (1996), Kaelbling et al. (1996), or Sutton and Barto (1998) for sound introductions to reinforcement learning). Agents take actions in the uncertain environment and obtain immediate rewards associated with these (and possibly previous) actions. A specific learning algorithm allows agents to adjust their action policies in pursuit of highest long-term rewards. It is a very desirable feature of any financial model that agents strive for strategic, as opposed to myopic, behaviour. This is exactly what reinforcement-learning agents do. On the other hand, it is the immense complexity of investors' interaction, both in real world financial markets and in the model that dramatically limits agents' abilities to actually achieve optimal investment policies or even makes the optimal investment behaviour outright impossible.

In our model we use a popular reinforcement learning algorithm, also known as the Q-learning, which was initially proposed by Watkins (1989). It is a temporal difference learning based on the step-wise update (or back-up) of the action-value function and associated adjustment of behavioural policies. The principal back-up rule is closely related to Bellman optimality property and takes the following form:

$$Q(s_t,a_t) \leftarrow (1-\alpha) \cdot \underbrace{Q(s_t,a_t)}_{\text{Old estimate of } Q(s_t,a_t)} + \underbrace{\alpha(r_{t+1} + \gamma \max_a Q(s_{t+1},a))}_{\text{New estimate of } Q(s_t,a_t)}. \tag{10}$$

Here s_t denotes the state of environment, a_t is the action taken in period *t* and r_{t+1} is the immediate reward associated with action a_t (and possibly earlier actions). Parameter α is known as the learning rate and γ is the discount rate of future rewards. Function $Q(s_t,a_t)$ is usually referred to as the action-value function (or Q-function) and it basically shows the value of taking action a_t in state s_t under behavioural policy π. More specifically, the action-value function is the expected cumulative reward conditional on the current state, action and pursued behavioural policy.

However, the so-called "curse of dimensionality" implies that a straightforward implementation of the basic version of this algorithm is rarely possible in complicated environments. Following the standard practice, we apply the Q-learning algorithm with gradient-descent approximation. Here we only describe specific variables that are used in the Q-learning algorithm.

As was mentioned before, there are two instances of individual agent learning in the model: learning to forecast dividends and learning to adjust perceived fundamentals. In the dividend forecasting case agent *i* learns to adjust the dividend adjustment factor, $a_{i,t}^{div}$ (see equation (2)). In each state there are three possible actions, i.e. the agent can increase the dividend adjustment factor by a small proportion specified by the modeller, decrease it by the same amount or leave it unchanged.

Due to the complex nature of environment, the state of the world, as perceived by investor *i,* must be approximated, and it is described by a vector of the so-called state features, $\vec{\phi}_s$. We choose four state features that are indicative of the reinforcement learner's "location" in the environment and summarize some properties of the dividend-generating process, which can provide a basis for successful forecasting. These features include the size of the dividend adjustment factor, relative deviation of current dividend from its EWMA (compared to the standard deviation), the square of this deviation (to allow for nonlinear relation with forecasts) and the size of the current dividend relative to the EWMA.

The forecast decision is taken at time *y* and the actual dividend realisation is known at forecast horizon *y* + *n*. Then agent *i* gets the reward, which is the negative of the squared forecast error:

$$r_{i,y+n}^{d} = -\left(d_{y+n} - E_y(d_{i,y+n})\right)^2. \tag{11}$$

Hence, the agent is punished for forecasting errors. The learning process is augmented by modeller-imposed constraints on dividend forecasts. The forecast is not allowed to deviate by more than a prespecified threshold (e.g. 30%) from the current level of dividends. In that case, the agent gets extra-punishment and the dividend forecast is forced to be marginally closer to the current dividend level. Once the agent observes the resultant state, i.e. the actual dividend realisation, it updates its behavioural policy according to the Q-learning procedure.

In the case of individual stock value estimation, agent *i* also can take one of three actions: fractionally increase or decrease the price adjustment factor, $a_{i,t}^{p}$ (see equation (5)), or leave it unchanged. Analogously to the dividend forecasting case, the four state features are the price adjustment factor, the stock price deviation from its exponential time-average (this difference is divided by the standard deviation), the square of this deviation and the current stock price divided by the weighted time-average.

The agent observes the state of the world and acts according to the pursued policy. After the trading round, the agent observes trading results and the resultant state of the world, which enables the agent to update its policies according to the usual Q-learning procedure. In this model, the basic immediate reward, $r_{i,t+1}^{p}$, is simply the log-return on the agent's portfolio:

$$r_{i,t+1}^{p} = \ln\left(h_{i,t}^{1} p_t + m_{i,t}^{1}(1+\bar{r}^{monthly})\right) - \ln\left(h_{i,t}^{0} p_{t-1} + m_{i,t}^{0}\right). \tag{12}$$

Recall that p_t denotes the market price following a trading round in time *t* and $\bar{r}^{monthly}$ is a one-period return on bank account. In order to ensure more efficient learning – just like in the case of dividend learning – constraints are imposed on the magnitude of price adjustment factors, and additional penalties are invoked if these constraints become binding.

The chosen specification of the reward function implies that the reinforcement-learning agents try to learn to organise their behaviour in order to maximise long-term returns on their investments. We could interpret agents in this model as professional fund managers that care about maximising clients' wealth, seek best long-term performance among peers and shun under-performance.

The model allows for optional alteration of agent behaviour via sharing private trading experience, competitive evolutionary selection and noise trading behaviour. These options help enhance the realism of the artificial stock market and arguably augment the reinforcement learning procedure by removing clearly dominated trading policies implemented by individual agents and by strengthening competition among them.

3. Experimental Simulations

Like the majority of other ASM models, the model presented in this paper is based on a large number of parameters, and it is very difficult to calibrate the model to match empirical data. At this stage of the model development we do not attempt to do that. Instead, we assign reasonable and, where possible, conventional values to the parameters and assume very simple forms of dividend-generating processes. This enables us to determine the approximate fundamental stock value dynamics and study how the market stock price, determined by the complex system of interacting heterogeneous agents, fares in relation to stock price fundamentals. Even though the model is not calibrated to the market data, the results provide qualitative insights into market self-regulation, efficiency and other aspects of market functioning. In this section we examine these issues in more detail and report some interesting simulation results.

The simulation procedure is implemented by performing batches of model runs. Each run consists of 20,000 trading rounds (about 1667 years). Batches of ten runs repeated under identical parameter settings are used to generate essential data and statistics that are, in turn, used for analysis and generalisation. In every run, the first 5,000 trading rounds – as the learning initiation phase – are excluded from the calculation of the descriptive statistics. The simulation concentrates on altering the features of the reinforcement learning, interaction among agents and dividend-generating processes in an attempt to understand relative importance of intelligent individual behaviour, market setting and population-level changes for the aggregate market behaviour. Other model parameters are kept unchanged.

We start with the examination of the agents' ability to forecast dividends. Since dividends are driven by very simple data generating processes, it is not surprising that in the model version in which both reinforcement learning and evolutionary selection are enabled, agents are able to form very precise forecasts. The average dividend forecast error for this model specification is -0.1%, while the average absolute forecast error amounts to 0.4%. To assess the actual importance of the reinforcement learning behaviour for dividend forecasting, simulation batches with disabled reinforcement learning are run. In these runs agents neither learn to forecast dividends, nor try to optimise their portfolios, as their commensurate reinforcement rewards $r^{d}_{i,t+n}$ and $r^{p}_{i,t+1}$ are set to zero. In this case, the average forecast bias considerably increases to -0.8% and the average absolute error stands at 1.4%. In this no-learning case the average percentage of agents hitting the modeller-imposed dividend forecast bounds increases significantly, as compared to the enabled learning case. In other words, learning agents are able to effectively form "reasonable" forecasts, while non-learning agents are simply forced to remain within the pre-specified boundaries but perform much worse, taken on an individual basis. This leads us to a very natural conclusion that in the dividend forecasting process intelligent adaptation matters.

As the next step of our analysis we examine the dynamics of the market price in relation to the fundamentals. In this experiment fundamentals anchor the stock price dynamics to some extent, and the market price fluctuates in the vicinity of the perceived fundamental value The average percentage bias of market price from the fundamentals is low and stands at -1.6%. Nevertheless, the valuation errors are clearly autocorrelated, i.e. due to the market inertia and prevailing expectations, the stock price may be above or below risk-neutral fundamentals for extensive periods of time. For instance, runs of uninterrupted overvaluation stretch on average for 44 trading periods and an average length of undervaluation runs is 60 periods. By the same token, average market price deviations from the fundamental valuation are large relative to the price volatility. The enabled evolutionary selection option in the model ensures even wealth distribution among agents and active agents in each trading period (i.e. agents that have sufficient funds and/or stock holdings to trade constitute on average 89.7% of total population). Finally, the average fraction of agents whose adjusted fundamental valuations (reservation prices) fall out of modeller-imposed "reasonable" bounds is very low and stands, on average, at 0.1% of total population in a trading round.

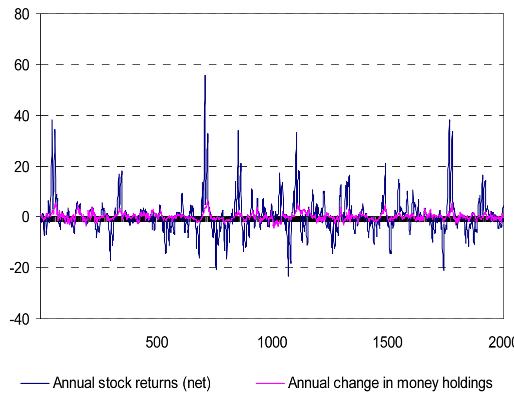
It turns out that the above results strongly depend on the evolutionary competition assumption. It suffices to disable the evolutionary selection, and the average percentage stock price bias from the fundamentals boosts to 5.9% along with a dramatic increase in average overvaluation runs to 406. By the end of a simulation run, the number of inactive agents per trading round increases to 70-80%, and wealth, naturally, concentrates in the hands of the remaining 20-30% agents. There are some possible explanations to this overvaluation and wealth concentration. Such overvaluation can be, to a certain extent, associated with the model's feature that excess liquidity is simply taken away from the market, which means that the agents that tend to sell their stock holdings are more likely to consume their money and become inactive. In other words, those agents that highly value the stock tend to dominate in the market. Another interpretation is that agents performing worse are simply driven out of the market. Moreover, a diminishing number of active participants and a smaller degree of competition allows agents to concert their portfolio rebalancing actions in such a way that the market price is driven up, which leads to larger unrealised returns and thereby stronger reinforcement for the remaining active players. These results make sense from the real world perspective. The largest mass of investors want stock prices to be as high as possible (though possibly still compatible with fundamentals), and it is not in their direct interest to have prices that match fundamentals precisely.

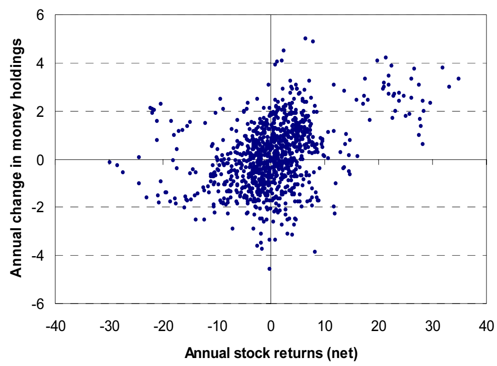
We also perform simulations to examine the market's self-regulation ability. In particular, we want to know whether economic forces are strong enough to bring the market to the true fundamentals if they systematically differ from average perceived fundamentals. For this purpose, we introduce an arbitrary upward bias to the estimates of the fundamental value by adding an arbitrary term in equation (3). Then simulation runs are implemented for different model settings, with or without reinforcement learning. It turns out that the market is not able to find the true risk-neutral fundamentals. In the case of no-learning, stock prices tend to slowly grow larger than the perceived fundamentals. In the case of enabled reinforcement learning, agents tend to stick to the perceived fundamentals, and, as a result, the market price fluctuates around them.

The above results confirm that the market self-regulation mechanism in this model is weak. We do not find evidence of agents adjusting their perceived fundamentals so that the market price gets in line with modeller-imposed fundamentals or, say, the usually assumed risk-averse behaviour. On the other hand, it is not surprising. Well known puzzles of empirical finance and recent mega-bubbles suggest that after all markets may not be tracking fundamentals so closely. It can be the case that markets exhibit an inertia so large that even fundamentally correct investment strategies pay out only in too distant future and may not be applied successfully or act as the market's self-regulating force. The obtained results suggest that (not necessarily objectively founded) market beliefs of what an asset is worth are a very important constituency of its market price.

Our last but not least intention was to examine the relationship between the market price fluctuations and the financial market liquidity. This experiment also helps to shed light on the reasons for a relatively loose connection between the market price and fundamentals. In this simulation run, the standard model version with reinforcement learning and evolutionary selection is used, while dividends are assumed to be deterministic and constant. It is notable that even in this environment market price fluctuations remain significant and trading does not stop. The clue to understanding this excess volatility may be the positive relationship between market liquidity and the stock price. Since unnecessary liquidity at an individual level is removed from the system, overall liquidity fluctuates in a haphazard way. Increases in market liquidity result in an increase in solvent demand for the stock and lifts its price. As can be seen from Figure 1, liquidity growth spikes are associated with strong price increases. The linear correlation between growth of money balances and stock price growth is found to be 0.32.

Figure 1. Typical relationship between stock returns and liquidityin a constant dividend case





It should be noted that the latter experiment is devised so as to ensure that positive relationship between stock returns and investors' cash holdings is not linked to fluctuations in dividend payouts. This allows us to conclude that liquidity fluctuations affect the asset price in this case, and not vice versa. The evidence that market liquidity changes can move markets is very important for understanding the way liquidity crises, credit booms and busts (deleveraging), portfolio reallocations between asset classes and other exogenous factors may affect stock markets.

4. A Short Presentation of Parallel Decisions Management System in Capital Markets.

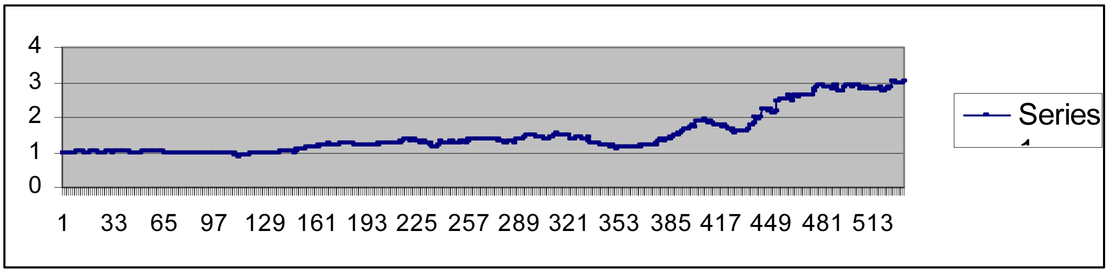
In this section we apply a parallel decisions management system in capital and exchange markets as an empirical counterpart of the so-called double trump model, which at first was designed for decisions management in exchange markets and later repeatedly used in various capital markets. The description of the double trump model, its development and possibilities of application for decisions management in exchange markets can be found in Rutkauskas (2005, 2006, 2008a, 2008b), Rutkauskas and Stasytytė (2006), Rutkauskas, Lukoševičius and Jakštas (2006), and Rutkauskas, Miečinskienė and Stasytytė (2008). This approach has close linkages to financial management research conducted by Ginevičius and Podvezko (2008a, 2008b), and Rutkauskas and Stankevičienė (2006). The analytical framework also benefits from conceptual ideas on modelling principles developed by Buračas (2004), Žvirblis, Mačerinskienė and Buračas (2008).

The link of parallel systems with the main topic of this paper and the ability of market participants and the market itself to match with consistent patterns of market behaviour and decisions management could be described by the following circumstances:

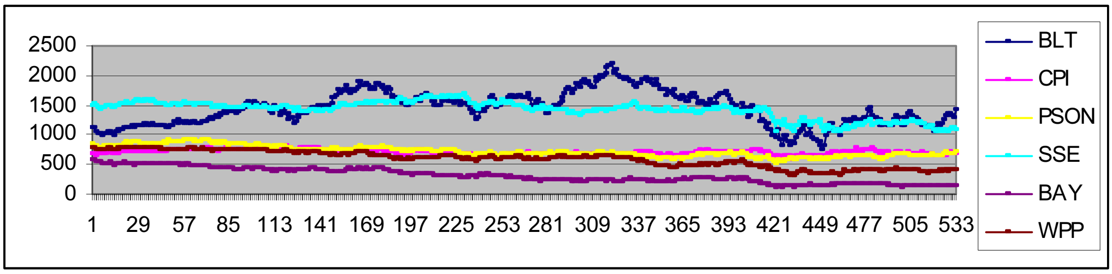
- under the circumstances of financial instability, capital markets in a sophisticated but cognizable manner change the supply of possibilities for an investor, which is fully described by possibilities' efficiency, risks and reliability;
- investor (either individual or institutional) has a possibility to perceive the adaptation principles and means of his utility function in changing behaviour of the capital market;
- perception of decisions management strategies and criteria interdependencies becomes a presumption and a guarantee of successful investing;
- decisions are made in an almost fully artificial market, which is compared with real market data only by the core parameters. The space of stochastic processes is an adequate enough real market model.

The possibilities of parallel decisions management system will be illustrated by its application for the achievement of the highest possible growth of invested capital during the analysed period: 02-01-07 – 09.04.09, which also includes the most severe periods of global financial crisis. In general, using this system a broad monitoring is maintained, which includes about 30 various global markets, for the search of favourable decisions for capital growth. The search of favourable decisions was performed selecting the so-called pseudo-scenario, when a part of historical data, in this case, 40 days from the beginning of 01-01-07, is accepted as "real" historical data, and the following data is treated as "forthcoming", and, with regard to the latter, forecasting is being performed and portfolio rebalancing decisions are being made. Next, the results of the experiment in five countries' markets will be presented: UK, Germany and the U.S. In sections *a* of Figures 2 – 4 we see how a unit of invested capital was changing during the analysed period, and in sections *b* we see the change of real prices of the 6 stocks in portfolio.

Figure 2. The results of the experiment in UK

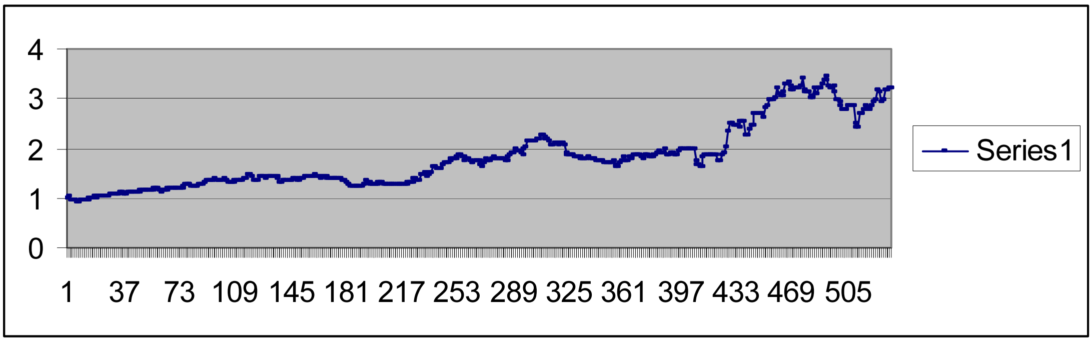


a) The change of a unit of invested capital

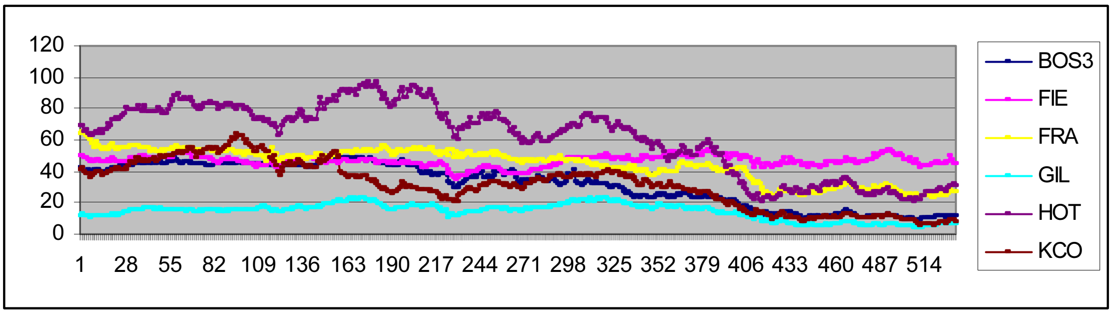


b) The change of real prices of stocks in portfolio

Figure 3. The results of the experiment in Germany

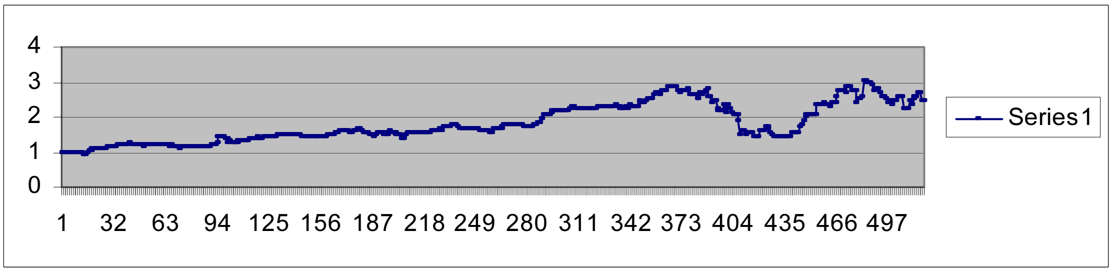


a) The change of a unit of invested capital

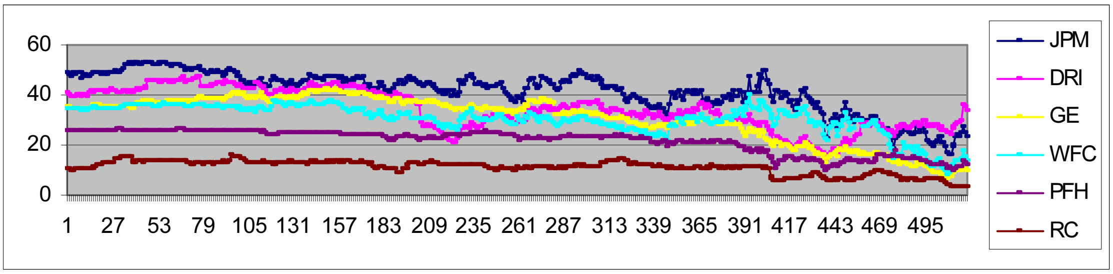


b) The change of real prices of stocks in portfolio

Figure 4. The results of the experiment in the U.S. (NYSE)



a) The change of a unit of invested capital



b) The change of real prices of stocks in portfolio

The growth of the initial invested capital is obtained with the help of the optimization (rebalancing) of portfolio structure. While rebalancing portfolio a fee of three basic percentile points was imposed in case of buying as well as selling a stock. This amounts to nearly 30% of the whole invested capital and more than 50% of the gross capital increase.

5. Concluding Remarks

In this paper we developed an artificial stock market model based on the interaction of heterogeneous agents whose forward-looking behaviour is driven by the reinforcement learning algorithm combined with an evolutionary selection mechanism and economic reasoning. Other notable features of the model include knowledge dissemination and agents' competition for survival, detailed modelling of the trading process, explicit formation of dividend expectations and estimates of fundamental value, computation of individual reservation prices and best order prices, etc. At this stage of development, the model should largely be seen as a thought experiment that proposes to study financial market processes in the light of complex interaction of artificial agents that are designed to act in an economically appealing way. Bearing in mind the uncertain nature of the model environment, mostly brought about by this same interaction, strategies followed by artificial agents seem to exhibit a good balance of economic rationale and optimisation attempts. Quite a strong emphasis on the model's economic content distinguishes this model from other ASM models, which are most often based on evolutionary selection procedures and are sometimes criticised for the lack of economic ground.

Preliminary simulation results suggest that the market price of the stock in this model broadly reflects fundamentals but over- or under-valuation runs are sustained for prolonged periods. Both individual adaptive behaviour and the population level adaptation (evolutionary selection in particular) are essential for ensuring any efficiency of the market. However, market self-regulation ability was found to be weak. The institutional setting alone, such as the centralised exchange based on the double auction trading, cannot ensure effective market functioning. Even in the case of active adaptive learning, the market does not correct itself from erroneously perceived fundamentals if they are in the vicinity of actual fundamentals, which underscores the importance of market participants' beliefs for the market price dynamics. We also found a positive relationship between stock returns and changes in liquidity, i.e. there are indications that exogenous shocks to investors' cash holdings lead to strong changes in the market price of the stock.

Parallel decisions search system, which is presented in the paper, exploits only a part of the market, i.e. a certain amount of stocks. The system is based on an assumption about market behaviour cognition, but the main model of market behaviour is admitted to be a multi-dimensional stochastic process, the identity of which regarding particular market is achieved with the help of stock prices, market indices and macro-economic data. The application of an expert system allows us to state that even under the circumstances of a global financial crisis distinct investment strategies are available, which guarantee long-term capital growth rates much higher than the general growth of the market.

References

1. Arthur, W. B.; Holland, J.; LeBaron, B.; Tayler, P. (1997). Asset Pricing under Endogenous Expectations in an Artificial Stock Market. In: Arthur, W. B.; Durlauf, S.; Lane, D. (eds). *The Economy as an Evolving Complex System II*, Addison-Wesley, Reading, MA, 15-44.
2. Bertsekas, D. P.; Tsitsiklis, J. N. (1996). *Neuro-Dynamic Programming*. Nashua: Athena Scientific.
3. Buračas, A. (2004). On Paradigm of Metaeconomics: Sense and Essense. In: *Management of Organizations: Systematic Research*, (29): 21–36.
4. Ginevičius, R.; Podvezko, V. (2008a). Multicriteria evaluation of Lithuanian banks from the perspective of their reliability for clients. In: *Journal of Business Economics and Management*, 9(4): 257–267.
5. Ginevičius, R.; Podvezko, V. (2008b). Daugiakriterinio vertinimo būdų suderinamumas. In: *Verslas: teorija ir praktika*, 9(1): 73–80.
6. Kaelbling, L. P.; Littman, M.L.; Moore, A. W. (1996). Reinforcement Learning: A Survey. In: *A Journal of Artificial Intelligence Research*, (4): 237-285.
7. Kim, G.; Markowitz, H. (1989). Investment Rules, Margin, and Market Volatility. In: *Journal of Portfolio Management*, (16): 45-52.
8. LeBaron, B. (2000). Evolution and Time Horizons in an Agent Based Stock Market. Brandeis University Working paper. Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=218309.
9. LeBaron, B. (2006). Agent-based Computational Finance. In: Tesfatsion, L.; Judd, K. L. (eds). *Handbook of Computational Economics*, Vol. 2, 1189-1233, Amsterdam: Elsevier.
10. Lettau M., (1997). Explaining the Facts with Adaptive Agents: The Case of Mutual Fund Flows. In: *Journal of Economic Dynamics and Control*, (21): 1117-1148.
11. Lux, T. (1995). Herd Behaviour, Bubbles and Crashes. In: *Economic Journal*, (105): 881-896.
12. Nenortaitė, J. (2007). A Particle Swarm Optimization Approach in the Construction of Decision-Making Model. In: *Information Technology and Control*, 36(1A): 158-163.
13. Nenortaitė, J.; Simutis, R. (2005). Adapting particle swarm optimization to stock markets. *Intelligent Systems Design and Applications: 5th International Conference, Wroclaw*, Conference proceedings, 520-525.
14. Plikynas, D. (2008). Portfolio Design and Optimization Using Neural Network Based Multiagent System of Investing Agents. *20th EURO Mini Conference "Continuous Optimization and Knowledge-Based Technologies", Neringa.*
15. Plikynas, D.; Simanauskas, L.; Būda, S. (2002). Research of Neural Network Methods for Com-

pound Stock Exchange Indices Analysis. In: *Informatica* 13(4): 465-484.

16. Ramanauskas, T. (2009). Agent-Based Financial Modelling: A Promising Alternative to the Standard Representative-Agent Approach. Bank of Lithuania Working Paper No. 3, Vilnius.
17. Rutkauskas, A. V. (2005). The Double-Trump Decision Management Model in Global Exchange. – In: *Ekonomika. Mokslo darbai*, (72): 84-104.
18. Rutkauskas, A. V. (2006). Adekvačiojo investavimo portfelio anatomija ir sprendimai panaudojant imitacines technologijas. In: *Ekonomika: mokslo darbai*, (75): 52-76.
19. Rutkauskas, A. V. (2008a). Investor's Possibilities Evaluation in Capital and Exchange Markets. *5th International Scientific Conference "Business and Management", Vilnius.*
20. Rutkauskas, A. V. (2008b). On the Sustainability of Regional Competitiveness Development Considering Risk. In: *Technological and Economic Development of Economy. Baltic Journal on Sustainability,* 14(1): 89-99.
21. Rutkauskas, A. V.; Lukoševičius, V.; Jakštas, V. (2006). Dvigubojo kozirio portfelio naudojimas sprendimams valdyti globalioje valiutų rinkoje. In: *Verslas: teorija ir praktika*, 7(2): 55–72.
22. Rutkauskas, A. V.; Miečinskienė, A.; Stasytytė, V. (2008). Investment decisions modelling along sustainable development concept on financial markets. In: *Technological and Economic Development of Economy: Baltic Journal on Sustainability*,14(3) 417–427.
23. Rutkauskas, A. V.; Stankevičienė, J. (2006). Integrated asset and liability portfolio as instrument of liquidity management in the commercial bank. In: *Journal of Business Economics and Management,* 7(2): 45–57.
24. Rutkauskas, A. V.; Stasytytė, V. (2006). The double trump portfolio as the core of sustainable decision making strategy in currency markets. - *The 10th World Multi-Conference on Systemics, Cybernetics and Informatics. July 16-19, 2006, Orlando, Florida,* 57-62.
25. Rutkauskas, A. V.; Stasytytė, V. (2008). Stratification of Stock Profitabilities – the Framework for Investors Possibilities Research in the Market. In: *Intelektinė ekonomika: mokslo darbų žurnalas*, 1(3): 65–72.
26. Samanidou, E.; Zschischang, E.; Stauffer, D.; Lux, T. (2007). Agent-based Models of Financial Markets. In: *Reports on Progress in Physics*, (70): 409-450.
27. Sethi, R.; Franke, R. (1995). Behavioural Heterogeneity under Evolutionary Pressure: Macroeconomic Implications of Costly Optimization. In: *Economic Journal*, (105): 583-600.
28. Simutis, R.; Masteika, S. (2004). Intelligent Stock Trading Systems Using Fuzzy-neural Networks and Evolutionary Programming Methods. In: *Solid State Phenomena*, (97-98): 59-64.
29. Stankevičius, G. (2001). Forming of the Investment Portfolio Using the Self-organizing Maps (SOM). In: *Informatica*, 12(4): 573-584.
30. Sutton, R. S.; Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge: MIT Press.
31. Tay, N. S. P.; Linn, S. C. (2001). Fuzzy Inductive Reasoning, Expectation Formation and the Behavior of Security Prices. In: *Journal of Economic Dynamics and Control*, (25): 321-362.
32. Watkins, C. J. C. H. (1989). *Learning from Delayed Rewards*. Ph.D Thesis, King's College.
33. Žvirblis, A.; Mačerinskienė, I.; Buračas, A. (2008). Įmonių konkurentų potencialo vertinimo principai ir baziniai modeliai. In: *Intelektinė ekonomika: mokslo darbų žurnalas.* 1(3): 82–91.

AGENTAIS PAGRĮSTAS AKCIJŲ RINKOS IMITACINIS MODELIS

Aleksandras Vytautas Rutkauskas
Tomas Ramanauskas
Vilniaus Gedimino technikos universitetas, Lietuva

Santrauka. Šiame straipsnyje pateikiamas dirbtinės akcijų rinkos modelis, pagrįstas heterogeninių agentų sąveika, kurią nulemia ekonominiai elgsenos principai, skatinamojo mokymosi algortimas bei evoliucinė agentų atranka. Šis imitacinis modelis vertintinas kaip struktūrizuotos analizės pagrindas, tiriant rinkos savireguliacijos galimybes, rinkos efektyvumą bei kylančias rinkos savybes lemiančius veiksnius. Lyginant su daugeliu kitų dirbtinės akcijų rinkos modelių, šiame modelyje ekonominei individų elgsenai ir individualiai adaptacijai skiriama gerokai daugiau dėmesio. Riboto racionalumo agentai šiame modelyje investicinius sprendimus grindžia ekonomine logika, t. y. vertindami tikėtinus diskontuotus pajamų srautus bei lygindami alternatyvių investicijų grąžas. Jie taip pat siekia tinkamai vertinti ateitį dideliu neapibrėžtumu pasižyminčioje aplinkoje bei atsižvelgia į kitų rinkos dalyvių veiksmų poveikį bendrai rinkos kainos dinamikai. Šis darbas yra vienas pirmųjų bandymų ekonominiu požiūriu įdomų skatinamojo mokymosi algoritmą (konkrečiau, Q-mokymąsi) dirbtinės akcijų rinkos modeliuose. Modelis taip pat pasižymi ganėtinai sofistikuota imitacinės rinkos struktūra.

Su modeliu atlikti imitaciniai eksperimentai, kurių metu buvo keičiami parametrai, lemiantys skatinamojo mokymosi, agentų tarpusavio sąveikos bei dividendus generuojančius procesus, siekiant įvertinti jų poveikį rinkos savireguliacijai, efektyvumui ir sisteminio lygmens dinamikai. Jų pagrindiniai rezultatai yra šie. Šiame modelyje akcijos rinkos kaina iš esmės atspindi rizikai neutralią fundamentaliąją vertę, tačiau galimi ilgi pervertinimo ir nepakankamo įvertinimo epizodai. Ir individualus mokymasis, ir populiacijos lygmens adaptacija yra esminės prielaidos imitacinės rinkos efektyvumui pasiekti. Rinkos savireguliacijos galimybės šiame modelyje yra silpnos. Imitacinėje rinkoje nustatytas teigiamas sąryšis tarp akcijų grąžos ir likvidumo (t. y. pinigų kiekio sistemoje) pokyčių.

Straipsnyje taip pat pateikiami preliminarūs bandymų taikyti analogiškus modeliavimo principus investicinio portfelio valdymui realiose finansų rinkose rezultatai. Jie patvirtina neblogas praktines šių modeliavimo principų taikymo perspektyvas.

Aleksandras Vytautas Rutkauskas is a Professor, Habilitated Doctor, the Dean of the Faculty of Business Management and the head of the Laboratory of Business Planning and Environment Economics at Vilnius Gediminas Technical University (Lithuania). Research interests: integrated value and risk management, regional development management, environment economics and real-time investment management in stock and exchange markets.

Aleksandras Vytautas Rutkauskas – Vilniaus Gedimino technikos universiteto Verslo vadybos fakulteto dekanas, Verslo projektavimo ir aplinkos ekonomikos mokslo laboratorijos vedėjas, profesorius, habil. dr. Pagrindinės mokslinių interesų kryptys: integruotas vertės ir rizikos valdymas, regionų plėtros valdymas, aplinkosaugos ekonomika, realaus laiko investicijų valdymas valiutų ir kapitalo rinkose.

Tomas Ramanauskas is a doctoral student at Vilnius Gediminas Technical University, Faculty of Business Management and the principal economist at the Economic Research Division of the Economics Department of the Bank of Lithuania. Research interests: financial market modelling, sustainable economic growth and macroeconomic and financial stability research.

Tomas Ramanauskas – Vilniaus Gedimino technikos universiteto Verslo vadybos fakulteto doktorantas ir Lietuvos banko Ekonomikos departamento Ekonominių tyrimų skyriaus vyriausiasis ekonomistas. Pagrindinės mokslinių interesų kryptys: finansų rinkų modeliavimas, tvarios ūkio plėtros, makroekonominio ir finansinio stabilumo tyrimai.