EMPIRICAL VERSION OF AN ARTIFICIAL STOCK MARKET MODEL

Tomas Ramanauskas

Lietuvos bankas
Gedimino pr. 6
LT-01103 Vilnius
E-mail: tramanauskas@lb.lt

This paper presents an artificial stock market model based on complex interaction of heterogeneous reinforcement-learning agents. The model is partly calibrated to empirical data and is aimed at analysing non-equilibrium market dynamics, examining market self-regulation abilities and other emergent properties, and enhancing generative understanding of recent episodes of financial booms and busts. Simulated market returns have realistic statistical properties, and the model offers some qualitative insights on recent actual financial market developments.
Keywords: agent-based financial modelling, artificial stock market, reinforcement learning, financial bubbles.

Introduction

In this paper we develop an artificial stock market (ASM) model based on complex interaction of heterogeneous reinforcement-learning agents. This is a refined version of our earlier model (Ramanauskas, Rutkauskas 2009), which is more parsimonious and better suited for empirical calibration. The aim of the proposed model is to examine non-equilibrium dynamics of a simple artificial stock market and calibrate it to the actual data. With the help of this model we want to examine market self-regulation abilities and drivers behind asset bubble formation. We compare properties of the simulated market dynamics to actual stock market returns and to stylised facts about financial market returns. We also aim at enhancing our generative understanding of recent episodes of financial booms and busts.

The proposed ASM model displays a reasonable balance of parsimony, specificity and realism. Like most other ASM models, this model is highly stylised – it has a simple market structure, the relatively low level of systemic complexity and largely *ad hoc* individual behaviour. Nevertheless individual agents are assumed to act sensibly from the real world investors' viewpoint. Agents try to maximise their long-term portfolio performance, they take into account both underlying stock fundamentals and the behaviour of others, and they try to adapt to changing and highly uncertain environment conditions. Agents' bounded rationality is determined by limited ability of behavioural algorithms to achieve optimal strategies due to high uncertainty and a complex interaction of agents. This contrasts with the modeller-imposed agent irrationality or deliberate neglect of certain determinants of stock value in some other agent-based financial models.

One of model's novel features is the application of the reinforcement learning algorithm* for governing agents' behaviour. Such behavioural principles are very interesting from the economic point of view, even though there are known caveats related to application of these algorithms in nonstationary, multi-agent environments. Unlike in many ASM models, model's realism is enhanced by partial calibration to actual risk-neutral fundamentals of stock prices. In addition to this, the model employs a realistic market price setting mechanism and is does not rely on any strong assumptions about the formation of aggregate supply or demand.

The paper is organised as follows. In Section 1 we provide a detailed description of model's building blocks and simulated market processes. Results of model simulation in the stationary external environment and assessments of market self-regulation ability

*More specifically, the Q-learning algorithm proposed by Watkins (1989).

■ Tomas Ramanauskas is a Senior Economist at Economic Research Division, Economics Department, Bank of Lithuania, and a doctoral student at Vilnius Gediminas Technical University, Faculty of Business Management, Department of Finance Engineering.
Areas of activity: financial market analysis, agent based modelling, macrofinancial stability analysis.

are provided in Section 2. We apply the model to the actual data in Section 3. This section includes a brief discussion about actual market developments and qualitative and quantitative comparison of simulated market dynamics to the actual market dynamics and to the stylised properties of financial markets. Since the current paper is an integral part of our broader research effort, we do not provide a review of the related literature but rather refer an interested reader to our earlier work (Ramanauskas 2008) for a review of related ASM models.

1. Description of the model

The ultimate goal of any model should be to enhance understanding of the nature of the real world processes. It does not suffice to achieve a superficial resemblance of statistical properties of model's generated aggregate time series to those observed in the real world. In an applied model one should not resort to vague metaphors of actual investor behaviour, and such methods of inquiry cannot be justified on pure grounds of empirical fit. It seems that proponents of agent-based financial modelling do sometimes risk being carried away with their ideas, such as artificial agents' (or strategies') "breeding" in evolutionary algorithms or "black box" decisions in simple artificial neural nets. At the other extreme, the standard neoclassical financial modelling also hinges on heroic assumptions of perfect rationality, agent homogeneity, rational expectations and Walrasian market clearing, which results in theoretical constructs so idealised that investment practitioners do not recognise themselves in those models. A great merit of agent-based modelling is that with the help of growing computing power researchers are enabled to deepen their understanding of complex economic systems via controlled simulation experiments. In order for this generative understanding to be real, strong emphasis should be put on the economic content of agent-based models – individual behaviour and market structure must be based on clear and economically sound principles. In this section we provide a detailed description of the proposed ASM model. More specifically, we discuss model's main building blocks, market setting and market processes.

1.1. Model's main building blocks

We start the presentation of our model with a cursory description of the basic market structure, agents and general market processes. As usual in ASM models, the stock market is very simple – only one dividend-paying stock (stock index) is traded in the market and its price is determined by exogenous fundamental factors and endogenous interaction of heterogeneous market participants. More specifically, there are an arbitrarily large number of heterogeneous learning artificial agents. Investors differ in their financial holdings, risk aversion, beliefs about the true stock value, etc. This ensures diverse investor behaviour even though the basic principles governing their adaptation are the same across population. Model's main processes are summarised in Table 1.

Table 1

Main building blocks of the empirical ASM model

Building block	Principal aspects of building block
Forming individual forecasts of corporate earnings	Econometric error-correction models of trend-reversion Individual sample size adjustment via evolutionary selection and pair-wise interaction
Estimating fundamentals and reservation prices	Discounted expected future corporate earnings (give individual estimates of risk-neutral fundamentals) Individual adjustment as a result of reinforcement learning (standard Q-learning with linear gradient-descent approximation)
Carrying out trades via the centralised exchange	Individual trading decisions (arrived at by comparison of individual reservation prices with the prevailing market price) Simultaneous submission of trade orders and random queuing of individual orders Double auction system

Source: formed by the author.

First, all agents apply simple econometric tools to obtain individual forecasts of corporate earnings and discount expected earnings by applying government bond rates of commensurate maturities to arrive at risk-neutral estimates of the stock value. The fundamental value estimates are then adjusted as a result of competitive interaction of agents and individual reinforcement learning. Thereby agents obtain individual reservation prices and make their trading decisions by comparing them with the prevailing market price. Agents place limit orders to buy or sell stocks on the centralised exchange, and trades are carried out in the double auction system. We are mostly interested in the resulting dynamics of the simulated market price and properties of market returns. The model uses some empirical data (is partly calibrated), which enables viewing simulation results in the context of actual market properties and some stylised facts about financial stock returns.

1.2. General market setting

In this partial equilibrium model there are basically two securities – a risky dividend-paying stock (stock index) traded on the daily basis and a locally risk-free short-term lending facility. An arbitrarily large number of adaptive agents engage in daily stock trading activities. If an agent has no exposure to the risky asset, then all of its funds are kept in a bank account, which pays an *a priori* known overnight rate. The interest rate is exogenously determined as a short rate of the government bond yield curve and changes daily together with the entire term structure of interest rates. The exogenously determined term structure is important in determining estimates of stock fundamentals. In order to keep the model tractable, investors are not allowed to trade government bonds.

In contrast to exogenous interest rates, the market price of the stock is determined by endogenous interaction of market participants that trade through the centralised stock exchange. One important constraint on the trading process is that agents are allowed to have either one share or zero exposure to the risky asset (this also requires the number of agents to be larger than the number of shares). This assumption greatly simplifies agents' decision making processes and makes agent adaptation easier. This constraint on individual holdings is justified on the grounds of our interest in systemic rather than individual behaviour: though each individual agent can supply or demand only one share, the aggregate supply and demand balance is much more flexible to vary. On the other hand, as agents are unable to accumulate many shares, all of them exert only marginal and roughly the same influence over the market price. As a result, market developments in the model are not driven by financial dominance of the most wealthy market participants.

1.3. Forming individual forecasts of corporate earnings

Expected company earnings and dividend payouts are among crucial determinants of the intrinsic stock value. Even though in standard models based on the efficient market hypothesis corporate earnings and dividend dynamics are not forecasted explicitly, it is usually implicitly assumed that some market players do conduct fundamental analysis, which ultimately gets reflected in stock prices. Hence, the fundamental analysis of earnings perspectives does matter. It is only that some theories are willing to go so far as to assume that communication among market participants is efficient enough for most investors not to bother inquiring into companies' financial books.

Here we propose the view that in the uncertain environment investors (*i*) form their individual beliefs about the risk-neutral value of a risky stock as some basic value anchor, (*ii*) acknowledge that the market price of the stock may fluctuate about or systemically differ from individual risk-neutral fundamentals due to various factors, such as investors' risk preferences, animal spirits or heterogeneity of beliefs, and (*iii*) flexibly determine their individual reservation prices in the process of adaptive interaction with the environment. The inertia of beliefs about future prospects, as well as the entirety of individual incentives and reward structures then determine market's aggregate attitude toward risk and, consequently, result in episodes of market euphoria or pessimism.

In the current model all agents form their individual forecasts of corporate earnings dynamics. We also allow for possibility to improve a given agent's forecasting model by replicating the more successful individual's model via a pair-wise competitive selection. The whole procedure can be described as follows.

Agents observe quarterly realisations of nominal corporate earnings generated by an exogenous, potentially nonstationary data generating process. For the pricing of a security, dozens of years of expected corporate earnings are potentially important, and no financial or economic model is capable of providing such forecasts with any accuracy. Hence, it is not possible to apply standard forecasting tools in such case. Furthermore, want to keep the forecasting procedure as simple as possible. For these reasons agents assume that aggregate earnings follow simple linear trends (that could in principle be related to technological progress and inflation) with cyclical and random fluctuation around them. Hence, agents simply identify linear time trends of the observed time series and its medium-term trend-reversion from the following basic long-run (1) and short-run (2) regressions:

$$y_q = \beta_{1,i}^{LR} + \beta_{2,i}^{LR} \cdot q + \varepsilon_{i,q}^{LR}, \tag{1}$$

$$\Delta y_q = \beta_{1,i}^{SR} \Delta y_{q-1} + \beta_{2,i}^{SR} \varepsilon_{i,q-1}^{LR} + \varepsilon_{i,q}^{SR}. \tag{2}$$

Here y_q denotes net corporate earnings in quarter q; β and ε denote, respectively, the ordinary least squares regression coefficients and residuals from regressions set up by agent i. As usual, Δ is the temporal difference operator. All agents apply the same model specification but, importantly, they apply the model to individual data samples of different sizes. In other words, in the artificial stock market there is a wide variety of beliefs about what – short-, medium- or long-term past developments – can provide an appropriate indication of future tendencies of corporate earnings. The idea is close in spirit to LeBaron's (2003) model, in which agents use different amounts of past data in deciding on their optimal trading strategies.

As it is obvious from the above equations, individual agents do not conduct rigorous econometric analysis, and we should not expect a typical real world financial trader to do that either. Rather, artificial agents are interested in (subjectively) identifying general trends prevailing in the market in a given period. Of course, such behaviour would not be of much economic interest if the amount of data that individual agents use were predetermined. For this reason we allow for some kind of simple evolutionary selection of market beliefs, which occurs through imitation learning*. Imitation learning at systemic level can occur if individuals are able to observe and copy other people's actions if they lead to desired outcomes. In economic imitation models agents may imitate an individual that is located next to the observer (Eshel et al. 1998), a randomly chosen individual (Duffy, Feltovich 2000), an individual with highest known utility (Kirchkamp 2000).

In this model imitation is designed as follows. Initially, each agent tries to identify the corporate earnings trend from a randomly assigned amount of data. As new quarterly data get available, agents rerun their regression equations, and in order to do that they have to decide what "data window" to use. For this purpose all agents are randomly matched in pairs. Matched agents then compare their last quarter's performances** (based on their beliefs about fundamentals) and in the ensuing quarter the worse performing agent either (*i*) with some predefined probability copies the window width parameter from its counterpart (i.e. has identical beliefs about future fundamentals) or (*ii*) randomly experiments with the sample size. The probabilistic procedure of strategy selection is very common in evolutionary algorithms, and it is employed in order to ensure that the system does not get stuck in local extrema. It is also undoubtedly desirable to see persistent diversity of beliefs about stock fundamentals in the highly uncertain environment.

There can be countless alternatives to the chosen competitive selection procedure. We could choose from many different criteria for the sample size selection, e.g. regression fit statistics, individual regression residuals of the preceding period, etc. In our case

*In the psychology literature it is also known as a observational learning (Brenner 2006).

**See how agent performance is defined in Section 1.5.

experimentation with alternative selection procedures did not have a substantial qualitative impact on the overall results. The main reason for our choice of the current agent performance as a criterion for individual model selection is that it could help explain the apparent discrepancy between perceived and actual fundamentals during bubble formation periods and shed more light on stock market bubbles and busts. For example, in periods of strong upward market movements financial media naturally tends to assign more weight to the commentaries of those investment professionals that are known to be more bullish (i.e. optimistic) and have bet on the market rise. Once market trends change, bearish (i.e. pessimistic) commentaries tend to dominate financial coverage. In other words, the market tends to listen to those that are "more successful" at the moment. More generally, a largely unexpected market change, if sustained for some period, can be identified by market participants as a qualitatively new market trend and it can become a self-fulfilling prophecy. In this fashion a significant temporary increase in corporate earnings (accompanied by a market tale such as a "technological breakthrough", "internet economy" or "housing boom") can create a strong precondition for a financial market bubble, as people tend to think that past tendencies are no longer valid due to a structural break. The specific design of the competitive selection procedure in our model is aimed at capturing such situations.

To form forecasts, each agent simply applies its individual model set out in equations (1) and (2). However, we have to put additional constraints on expected corporate earnings in order to avoid unsustainably steep upward trends or negative values in the long term. Hence we assume, quite naturally, that by using the above models agents form only medium-term forecasts (e.g. 10 quarters ahead forecasts). Due to huge intrinsic uncertainty surrounding any long-term economic predictions, it is further assumed that beyond the forecast horizon expected corporate earnings flatten out and equal the agent's expected average value of earnings within the forecast horizon. Negative long-term corporate earnings expectations are also ruled out by setting them equal to 0.

1.4. Estimating fundamentals and reservation prices

In this subsection we describe how agents get to their trading decisions. Individual expectations of the future corporate earnings path are relevant as primary inputs in the calculation of individual estimates of the fundamental stock value. As agents interact with the environment, they adjust their valuations to obtain individual reservation prices. Trading decisions are determined by a straightforward comparison of the prevailing market price with the individual reservation price.

Let us first clarify the notion of the fundamental stock value. In this context the fundamental stock value is interpreted as the risk-neutral valuation of expected future earnings:

$$v_{i,q}^{fund} = \frac{y_{i,q+1}^{e}}{1+r_{q,q+1}^{f}} + \ldots + \frac{y_{i,q+n}^{e}}{(1+r_{q,q+n}^{f})^{n}}.$$

In this equation $v_{i,q}^{fund}$ denotes agent i's estimate of fundamental value in quarter q. Agent i's individual forecast of corporate earnings j quarters ahead is denoted by $y_{i,q+j}^{e}$. There are n terms in the equation and n is sufficiently large to ignore the rest of the infinite stream of expected earnings. Discount factors are based on government bond yields of appropriate maturities, $r_{q,q+j}^{f}$. The term structure of interest rates varies daily and so do individual estimates of the fundamental value. Discount rates are adjusted to approximately reflect the actual remaining time to earnings realisation.

The risk-neutral discounting and agents' attitude to risk deserve some commentary. First of all, the above estimation of fundamentals does not imply risk-neutrality of agents as it is only an interim calculation which is further adjusted for risk. However, for several reasons we want to avoid imposing the standard hard-wired risk aversion in agents' utility functions. One reason is that changes in investors' collective attitude to risk, e.g. panic or euphoria, are often an important driver behind dramatic market movements unjustified by fundamentals. Another technical reason is that there is no consumption in this model, so risk aversion cannot be linked to consumption smoothing behaviour.

Agents in this model should be interpreted as institutional investors, e.g. professional asset managers, and their attitude to risk can substantially differ from that of individual investors. Rather than concentrate on consumption smoothing, professional asset managers in principle should care about maximising clients' wealth, seek best long-term performance among peers and shun under-performance. One would expect a professional asset manager to be considerably less risk-averse than a prudent individual. This is because institutional investors have the expertise needed to act in the risky environment, their reward schemes are known to often be asymmetric (due to sharing profits but not losses), the riskier investment strategies are associated with higher average returns and thereby larger asset management fees, etc. Moreover, institutional investors and, more generally, sophisticated market players often engage in what is known as "trend riding" (Frankel, Froot 1987). They may buy an asset, which strongly rises in value, even though they may consider it overvalued. Buying in the early stage of the bubble formation still leaves quite a lot of room to square the position with considerable profit once the trend reverses. In more technical terms, periods of bubble formation skew return distribution – the probability mass concentrates at higher returns but the probability of catastrophic outcomes rises. Catastrophic outcomes are by definition rare and their cost is mostly borne by the clients of asset managers due to the above mentioned reward scheme asymmetry. This may provide some explanation for the mass euphoria among institutional investors during bubble formation periods. At the same time it is very difficult to explain these phenomena by resorting to utility-based risk measures and even more so by the consumption smoothing paradigm.

There are obvious problems with standard measures of risk aversion, so we take a different approach to measuring investors' attitude to risk. An investor is risk averse if he is willing to sell the asset at a lower price than his perceived fundamental value (defined as above). Conversely, he is risk loving if he is willing to buy the asset at a price above his perceived fundamental value. We refer to this perceived value of the risky asset as the individual reservation price because an investor wants to buy (sell) the asset when the individual reservation price is above (below) the observed market price.

In trading period t agent i's reservation price $v_{i,t}^{reserve}$ is assumed to be determined by three elements – a change in perceived fundamentals, prevailing market price and the reinforcement learning adjustment factor:

$$v_{i,t}^{reserve} = \frac{v_{i,t}^{fund}}{v_{i,t-1}^{fund}} \cdot p_{t-1} \cdot (1 + a_{i,t}). \tag{3}$$

The logic behind this pricing equation is simple. Each agent has its individual perceptions about changes in "raw" fundamentals. Prevailing market price cannot be ignored either because it reflects the consensus of all market participants. Every agent also makes some further subjective adjustments in an attempt to find the "right" reservation price leading to best investment decisions. Adjustment factor $a_{i,t}$ is chosen through the individual learning process, which takes place while the agent interacts with the environment. Note that in this setting the individual attitude to risk depends on attitudes of other agents. This is also related to the below discussed assumption that individuals want to achieve good performance relative to others. The market price determined by the entirety of individual reservation prices might systemically differ from perceived fundamentals, and it is reasonable for an individual investor to take into account this market consensus if he wants to achieve competitive performance. For instance, in a protracted period of bubble formation, a cautious asset manager choosing not to invest in the risky asset would systemically underperform relative to others (and possibly would even be forced out of the market), which induces clear incentives to take risks, i.e. be less risk averse.

1.5. Agents' reinforcement learning

Let us now turn to the learning process through which individual agents' pricing considerations, attitudes to risk and, more generally, goal-oriented behaviours are determined. Quite some learning methods are known, ranging from psychology-based models (stimulus-response, belief-based conscious learning, associative learning, etc.) to rationality-based methods (Beyesian, least-squares learning) to artificial intelligence approaches (evolutionary algorithms, replicator dynamics, neural nets, reinforcement learning)*. As Brenner notes, virtually all of the learning models used in economic contexts are largely *ad hoc*, based only on introspection, common sense, artificial intelligence research or psychological findings.

In this model we choose the reinforcement learning technique, borrowed from the machine learning literature. Reinforcement learning algorithms are rarely used in economic problems and to our knowledge have not been used in the artificial stock market models but they seem to be conceptually suitable for modelling investment behaviour**. Our choice of the learning algorithm is largely determined by the economic appeal of the reinforcement learning behaviour: by taking actions in the environment and obtaining immediate rewards associated with those or previous actions, a reinforcement-learning agent tries to find optimal policies, which maximise long-term rewards. Similarly, in standard economics and finance an economic agent often chooses investment allocation, consumption or other action plans so as to maximise the life-time utility. These two problems are conceptually very similar, and the main difference is that a reinforcement-learning agent, realistically, does not have to know the underlying model of reality and tries to infer world state transition probabilities by gathering experience from the interaction with environment. Importantly, the reinforcement learning algorithm, unlike many simple forms of adaptive learning, is well-positioned to solve the temporal credit assignment problem, i.e. in theory it is capable of determining strategic actions that enable agents to reach their strategic goals even though those actions may not give desirable results in the short-term. We certainly want agents in our model to strive for strategic, as opposed to myopic, behaviour. 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 if not makes the optimal investment behaviour outright impossible.

The specific learning algorithm used in this model is known as the Q-learning and was initially proposed by Watkins (1989). It is the 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:

$$Q(s_t, a_t) \leftarrow (1-\alpha)\cdot \underbrace{Q(s_t, a_t)}_{\text{Old estimate of } Q(s_t, a_t)} + \underbrace{\alpha(r_{t+1} + \gamma \max_a Q(s_{t+1}, a))}_{\text{New estimate of } Q(s_t, a_t)}. \tag{4}$$

Here s_t denotes the state of environment, a_t is the action taken in period t and r_{i+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 the 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:

$$Q^{\pi}(s,a) = E_{\pi}\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s, a_t = a\right\}.$$

So the back-up rule (4) governs the process of updating the unknown value of an action taken in a given state of environment. The intuitive interpretation of the rule is as follows: (*i*) the agent takes an action, which moves it from state s_t to s_{t+1} and triggers immediate reward r_{t+1}, then (*ii*) the agent evaluates how favourable is the resulting state with regard to its strategic goals, provided that it will further pursue strategic behaviour (which simply means that the agent sweeps through all possible actions in state s_{t+1} to find the highest estimate of the action-value function), and finally (*iii*) the agent looks back on the previous state-action pair and updates its value as the weighted average of

*For an overview of popular learning algorithms see, e.g., Brenner (2006).

**See Ramanauskas (2008) for discussion.

***For a presentation of principles of reinforcement learning algorithms see, for instance, Sutton and Barto (1998), Bertsekas and Tsitsiklis (1996) or Mitchell (1997).

the old and the new estimate. Importantly, it has been shown under quite general conditions that if the agent follows the so-called ε-greedy action, i.e. most of the time chooses actions associated with highest Q-values and occasionally takes exploratory (possibly random) actions, and if it visits all state-action pairs sufficiently many times, the update rule (4) guarantees convergence of the action value function to the optimum and consequently guarantees finding the optimal strategy.

Yet there is a practical implementation problem with the Q-learning algorithm. It can be used in its simple form only if the state-action space is relatively small so that a look-up table of the Q-function values could be formed. In most real world problems the sheer size of the Q-table and the computational burden associated with back-up operations are unmanageable (it is the so-called curse of dimensionality problem). This necessitates generalisation of agent experience, which can be achieved by approximation of the action-value function and discretisation of possible actions. In the model we apply a standard linear gradient-descent function approximation for the Q-function (see Ramanauskas, Rutkauskas 2009).

We now describe the specific variables that are plugged into the general back-up rule. At time t agent i takes action $a_{i,t}$ to incrementally adjust its individual reservation price in equation (3). The set of possible actions is small but it is not uniform in the sense that in the adaptation process an agent is allowed to make adjustments of different sizes. The state features that characterise the state of the environment are basically technical indicators which carry the essential information needed for the assessment of the adequacy of individual reservation prices. More specifically, agent i's state features vector consists of (*i*) one-year bond rate, (*ii*) individual perception of risk-neutral fundamental value, (*iii*) prevailing market price, (*iv*) deviation of the market price from the exponentially weighted linear trend, (*v*) deviation of the market price from its exponentially weighted moving average, and (*vi*) time remaining to the next earnings announcement.

We link immediate rewards to agents' net returns on wealth and also allow for the optional loss aversion that is to be discussed below. In this context it is useful to elaborate a little on the time line of intra-day events. It is assumed that trading may occur at the beginning of each trading day. If an agent successfully sells the asset, the funds get immediately transferred to the bank account and earn interest at the end of the day. Conversely, if an agent buys the asset, the money is deduced from its bank account and does not earn any riskfree interest but the owner of the stock is entitled to receive the dividend, which is assumed to be paid out at the end of day. Since quarterly dividends are assumed, non-zero dividend payments will of course be paid only once per quarter.

Hence, the net return on wealth are calculated as follows:

$$m_{i,t}^{1} = (m_{i,t}^{0} - b_{i,t} \cdot p_{i,t}) \cdot (1 + r_{t,t+1}^{f}) + h_{i,t}^{1} \cdot d_t ,$$

$$w_{i,t}^{1} = h_{i,t}^{1} \cdot p_t + m_{i,t}^{1} ,$$

$$r_{i,t}^{return} = w_{i,t}^{1} / w_{i,t-1}^{1} - 1 .$$

Here $m_{i,t}^{0}$ and $m_{i,t}^{1}$ denote agent i's money holdings before and after a trading round in time t. Similarly, $h_{i,t}^{1}$ indicates individual stock holdings at the end of a trading round, and this variable is restricted to take value of either 0 or 1. Indicator variable $b_{i,t}$ takes value of 1 if agent i buys the stock, it is equal to –1 if it sells the stock and equals zero otherwise. Agent i's actual traded price is denoted by $p_{i,t}$, whereas the average traded price, or the market price, is given by p_t. Variable d_t denotes the dividend payout, the same for all stockholders.

The return on wealth might not properly perform its function as a reinforcement signal if the money stock increases without bounds. In such case the impact of investment decisions on agent's wealth would gradually decline. Another possibility is that excess liquidity might create upward pressures on the price of the risky asset. For these reasons excess liquidity is removed from the system. In other words, the money that cannot be efficiently used for investment purposes is simply consumed.

From the reinforcement learning perspective, in any given trading period agents perform one full iteration step of strategy adaptation. At the beginning of the trading period they observe the state of environment (state features associated with last period's prices), take actions (adjust perception of fundamentals), observe the new state (state features at the end of the trading round), obtain immediate rewards (associated with standardised returns) and adapt their investment strategies. So despite the different timing notation, individual reinforcement signal $r_{i,t+1}$ can be directly calculated from standardised returns $r_{i,t}^{return}$, as the timing is actually the same.

To attain actual reinforcement signals, we make an important assumption that agents care not only about their returns but also about their relative performance. In line with our earlier discussion about agents' as professional asset managers' attitude to risk, we assume that one agent that has worst returns in a given period is punished by imposing an arbitrary negative reward. These additional costs could be associated with client loss or damaged reputation. This setting might in principle support bubble formation because in the case of an upward-trending market agents that invest in the riskfree asset would underperform more severely and bear greater risks of punishment, which would in turn add to incentives to invest in the risky asset.

A couple of other options for modelling reinforcement signals were also implemented in the model. As one option, negative returns are multiplied by a constant larger than one to reflect investors' possibly asymmetric treatment of profits and losses. This corresponds to the prospect theory developed by Kahneman and Tversky (1979), which *inter alia* states that people tend to strongly prefer avoiding losses to acquiring gains. Another possibility is to augment competitive co-evolution, e.g. by using relative returns (e.g. scaled by market returns) as the reinforcement signal. This rests on the idea, sometimes referred to as the Red Queen principle, of encouraging relative performance to spur systemic adaptation (see Martinez-Jaramillo 2007). It is natural that professional asset managers shun worse-than-average performance.

1.6. Carrying out trades via the centralised exchange

Having formed their individual beliefs about the fundamental value of the stock price, agents have to make specific portfolio rebalancing decisions. In principle, they weigh their own assessment of the stock value against the prevailing market price and make orders to buy (sell) one unit of the underpriced (overpriced) stock:

$$\textit{If } v_{i,t}^{reserve} > p_{t-1} \textit{ and } h_{i,t}^{0} = 0 \textit{ and } m_{i,t}^{0} \geq v_{i,t}^{reserve} \rightarrow \textit{ make order to buy 1 share,}$$

$$\textit{if } v_{i,t}^{reserve} < p_{t-1} \textit{ and } h_{i,t}^{0} = 1 \rightarrow \textit{ make order to sell 1 share, otherwise, make no order.}$$

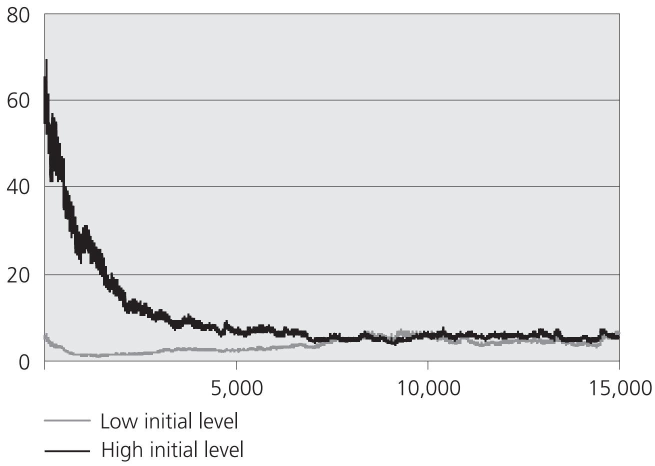
In the model agents interact via the centralised exchange by placing competitive limit orders, i.e. orders to trade the security at a specified or better price. Each agent is simply assumed to set the bid (ask) price to its reservation price. The trading mechanism is designed as the double auction system (more specifically, a simplified order book), in which both buyers and sellers submit their competitive orders to trade the stock. One, very common in other ASM models and uncontroversial assumption is that all orders are submitted at the same time without anyone knowing actions of others. Following the trading round, all agents' cash and securities accounts are updated accordingly. The centralised stock exchange also produces a number of trading statistics, some of which serve as a further input in agents' decision processes and are important systemic variables for the analysis of model results. For instance, 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. Simulation in stationary environment and examination of market self-regulation ability

In this section we perform some simple simulation exercises to examine market's self-regulation abilities. The idea is to assume different initial price levels and examine, *ceteris* *paribus*, whether simulated market prices determined by complex interaction of agents can reach similar levels associated with the same underlying fundamentals. We basically want to know whether competitive reinforcement-learning behaviour in this market setting allows agents to collectively determine relative prices of two assets (the stock and the instantaneously riskfree bond) in some systematic way.

To examine system's behaviour we assume stationary exogenous processes and perform model simulations for different parameter settings and different initial values of variables (see Tables 1 and 2 in the Appendix). In particular, quarterly corporate earnings are assumed to fluctuate randomly around a steady level, and a constant fraction of earnings is paid out as quarterly dividends. Interest rates are also assumed to be stationary and exhibit only minor variation in time. Simulation experiments are implemented with different values of some important variables or parameters, namely, the initial market price, the underperformance penalty rate and the loss aversion parameter (see Figure 1).

Figure 1. Simulated market price dynamics from different initial price levels (stationary external environment; penalty rate = –1; no loss aversion)



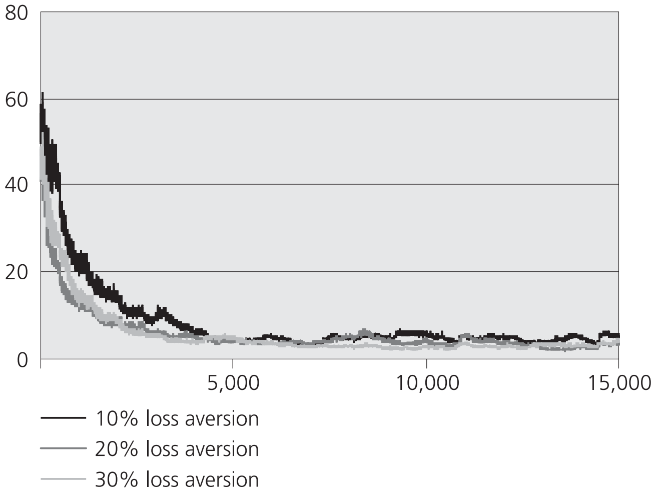
Source: model simulations.

It is obvious from the Figure 1 that simulation runs with different initial market prices, *ceteris paribus*, tend to converge to each other and exhibit similar movements. It is also interesting to note that given the same fundamentals, the remaining differences in price dynamics clearly indicate that market efficiency in the form of perfect congruence between market price and fundamentals is not possible because the market always retains some random element. Nevertheless the said convergence of price dynamics can be interpreted as some evidence for market self-organisation.

We can assess the importance of competitive behaviour for model results. Alteration of the penalty rate for the worst-performing agent did not have any significant impact on the equilibrium level of the market price. On the other hand, there is some evidence that stronger competition determined by larger penalty rates induces faster price convergence to the equilibrium in the high initial price cases. Moreover, if there is no punishment for the worst-performing agent and thus there is no direct competition among agents, the market prices decline to negligible levels. From the standard risk-reward perspective this results may seem somewhat counterintuitive, as one could expect agents' risk-neutral behaviour in this case. An explanation to this can be as follows. For all penalty rates the market price initially tends to fall because the riskfree asset offers immediate returns whereas stock dividends are paid only infrequently. Keeping the stock, which rapidly loses value, leads to wealth losses and reinforces negative price tendencies. Generally it takes time for agents to learn that there is some value in stocks, which becomes more evident as low stock prices result in large dividend rates and commensurately larger reinforcement rewards. However, in the case of non-competitive learning larger returns do not create sufficient pressures to buy the undervalued stock because stockholders already have it (so they are happy) and bondholders do not get negative reinforcement signals for their underperformance (i.e. they don't care). Also, imposing moderate penalties on worst-performing agents hinders the process of dumping the stock.

It is difficult to conduct any rigorous analysis of the simulated market's efficiency because there are no theoretical benchmarks for the proposed market setting. We can only note that the risk-neutral valuation of the discounted dividend flow in this stationary setting is close to 32, whereas the simulated market price converges the much lower price level of 4–7. The simulated stock price is about 3–4 times larger than annual dividends, which implies agents' large degree of risk aversion.

Figure 2. Simulated market dynamics for different loss aversion parameters (stationary external environment; penalty rate = –2)



Source: model simulations.

Rewarded risk taking can be naturally induced in the model by assuming agents' loss aversion. If they treat positive and negative stock returns asymmetrically, the risky asset becomes less attractive relative to the bond, which can only yield a positive return. As a result, due to lower demand the stock price gets smaller and dividend payouts become a larger fraction of the stock price, thereby implying larger stock returns. As can be seen from Figure 2, there is some evidence that larger loss aversion tends to lower stock prices and increase stock returns. For instance, if agents' reward functions are altered to increase negative returns by 10 per cent – the stock price fluctuates around 4.7, a 20 per cent increase gives the equilibrium price level of roughly 3.7, and 30 per cent asymmetry leads to price fluctuating around 3.1 in this setting.

3. Applying the model to the actual data

In this section we apply the model to the actual data in order to get some idea about empirical relevance of the simulated market dynamics and to compare properties of the generated market price dynamics with those of the actual market. Of course, we cannot expect to obtain a good match between simulated and actual price levels but the model does provide some qualitative insight on the boom and bust episodes in actual data. Important parameters describing the model and the experimental setup are given in Table 1 and Table 3 in the Appendix.

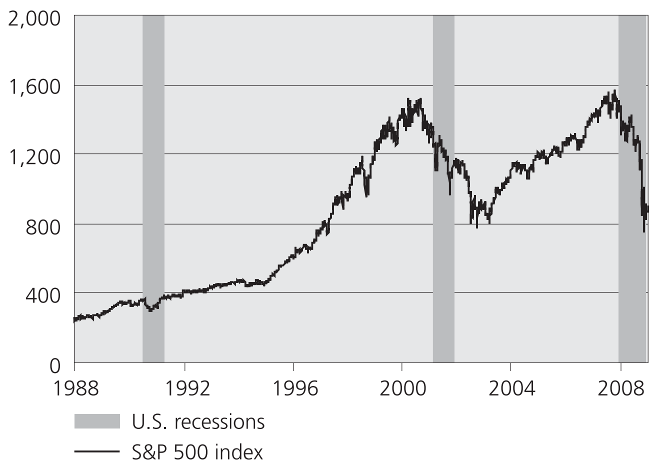
3.1. Model's data and background discussion

In this model the stock can be interpreted as a broad market index, which necessitates to have specific empirical data, such as aggregate earnings and dividends of companies included in the index and structural changes of the actual index composition. One of the few possibilities is to examine the broad U.S. market index S&P 500, for which the required data is provided by Standard and Poor's. Rapid financial integration of the last few decades and the recent global financial contagion make the examination of this specific index none the less interesting.

The data covers a 21-year period from 1988 to end-2008. The Standard and Poor's data set includes quarterly data on company earnings and dividends, and structural adjustment factors for the S&P 500 index. Also necessary for the simulation is the data on the term structure of interest rates. We use historical daily series of Treasury rates of constant maturities collected by the U.S. Federal Reserve, and apply a simple linear interpolation for all needed maturities for which data are not available. The series of daily S&P 500 index (with dividends excluded) is not used in calculations directly but rather serves as a comparison benchmark.

The actual data set includes two major boom-and-bust episodes. The first of these stock market booms occurred in the second half of the nineties. This boom episode was initially fostered by the increased productivity, output, employment, investment and wage growth (Jermann, Quadrini 2002). The rise of the revolutionary information technologies and favourable changes in financing conditions, buoyed venture capital investment in technology firms and contributed to the overall exuberance about the perceived structural shift to the "new economy". Around 1998, these developments seamlessly turned into the speculative bubble, later dubbed the "dot-com" bubble, which culminated in 2000. The bubble burst against the background of the faltering economy, worse than expected corporate earnings and the contractionary monetary policy pursued by the Federal Reserve. Of course, it was the technology companies' shares, included e.g. in the NASDAQ Composite index, that were mostly affected during this boom-and-bust episode but these developments were forceful enough to have impact on many broad indices across the world, including the S&P 500 (see Figure 3).

Figure 3. Dynamics of S&P 500 index and official timing of U.S. recessions



Source: Standard and Poor's and National Bureau of Economic Research data.

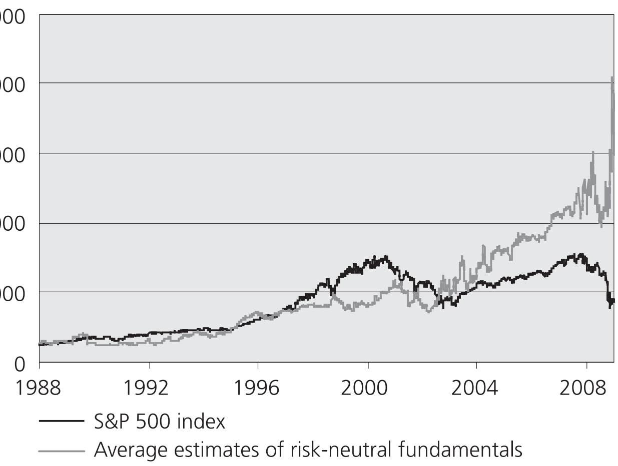
The U.S. stock market crash, accompanied by the economic recession of 2001 and accounting fraud scandals, had a negative impact on investors' confidence and their willingness to take risks in the stock market. As a result, the stock market performance remained subdued until 2003. Importantly, the Federal Reserve was fighting the 2001 recession by pursuing very aggressive monetary easing, which played a crucial role in house price bubble formation. Poor financial market performance also contributed to this, as excess liquidity and speculative capital made its way to the real estate market and consolidated housing as a means of lucrative investment. However, it did not take long before ultra-low interest rates, accessible financing and the real estate frenzy raised corporate earnings and once again greatly improved companies' medium-term prospects.

During this boom the S&P 500 doubled from its local trough dated October 9, 2002 to its all-time high close of 1,565 on October 9, 2007. Something that started with stagnation of the U.S. housing market, failures of highly leveraged market participants and the liquidity dry-up in the interbank markets soon turned into the biggest since the Great Depression global financial and economic crisis, marked by systemic bank failures, industry and even country bailouts, unprecedented global monetary and fiscal easing, and asset market crashes that wiped out trillions dollars of wealth. The crisis severely worsened in autumn 2008 with the collapse of Lehman Brothers. The S&P 500 index fell by 18 per cent in one week starting October 6. From peak to trough (recorded on March 9, 2009) the S&P 500 lost 57 per cent. Later, U.S. bank bailouts measured in trillions of dollars, huge fiscal stimulus and outright monetisation of the government debt contributed to a strong, though possibly not sustainable, stock market rebound from these lows.

3.2. Comparison of simulated and actual market dynamics

Now we qualitatively compare model results with the actual market dynamics. Recall that risk-neutral valuations of simple trend-reverting earnings projections serve in this model as some measure of stock fundamentals. It turns out that average agents' valuations exhibit very little variation across different simulation runs and do not systematically depend on risk parameters. Figure 4 compares fundamental valuations, averaged over 30 simulation runs and adjusted for the changing index structure, with the actual dynamics of the S&P 500 index. The findings are consistent with the earlier discussion. Stock prices are not justified by fundamentals during the technology boom, whereas in the second boom episode the estimated fundamentals could be associated with considerably higher stock prices. This suggests that the second stock market boom could be characterised by either heightened investors' risk aversion or perceived unsustainability of strong earnings growth and temporariness of the favourable interest rate environment.

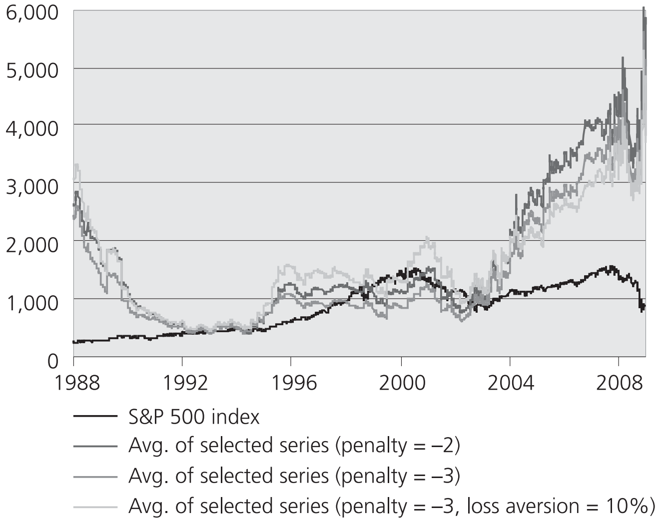
Figure 4. Average simulated estimate of fundamental value compared to S&P 500 index



Source: Standard and Poor's data; model simulations.

As a next step of our analysis, we perform batches of stock market simulation runs for different sets of risk parameters. Each of these batches consists of ten simulation runs from randomly chosen initial price levels. As was discussed in the previous section, simulations from low initial price levels often lead to even further price declines, so we remove these runs from the analysis. In Figure 5 we report simulated price paths averaged over economically interesting paths. We can see from the graph that simulated market prices decline from initial levels to fundamentals in about 4 years, and further generated series can be used for meaningful economic analysis. At this stage it is difficult to discern specific systematic features of price dynamics simulated with different risk parameters, as they look qualitatively very similar.

Figure 5. Simulated market prices compared to S&P 500 index



Source: Standard and Poor's data; model simulations.

Importantly, all generated series exhibit boom-and-bust behaviour. What is even more interesting is that the timing of the start of boom and bust periods closely corresponds to the actual developments. This can be explained by the observation that both artificial and real stock market's booms and busts coincide with significant changes in fundamentals. Also note that boom or bust episodes tend to begin more sharply in the model than in the real market, which may be attributed to agents' stronger herding behaviour in the model due to lower fundamental variety of agent expectations than in the real world. This suggests that the model may be useful for predicting rises and bursts of financial bubbles.

It should be noted that while in the first boom episode the simulated price level is broadly in line with the actual peak, during the second boom simulated prices strongly exceed actual prices. This gap can be partly explained by the fact that in the model earnings projections are simply based on past data and there was no way for agents to correctly anticipate the massive earnings correction that eventually materialised. Also, due to the lack of micro-evidence we did not calibrate or change agents' reward parameters to reflect possible changes in investor attitude to risk following the burst of the "dot-com" bubble.

3.3. Analysis of simulated returns properties

Like most ASM models, the current model is not directly intended for forecasting stock price levels. We are more interested in comparing properties of simulated stock returns with known stylised properties of stock returns and with actual properties of S&P 500 index returns. Some of the more important stylised empirical facts about stock returns can be summarised as follows (see Cont 2001):

1. *Non-normality of returns:* stock returns do not obey the Gaussian distribution, as their distribution tends to display heavy tails and asymmetry. However, this property wanes and the distribution becomes more Gaussian-like with the increasing time scale.

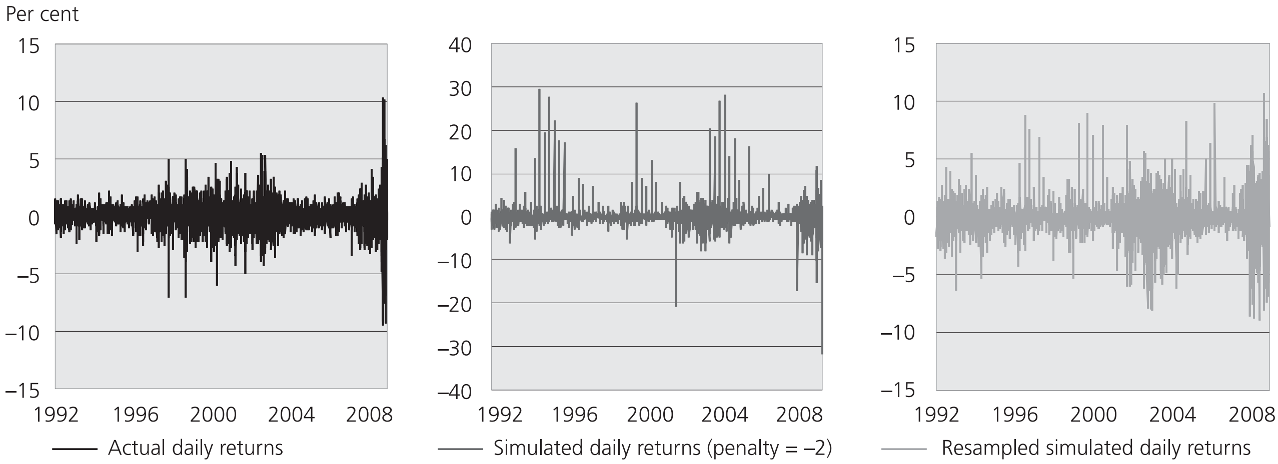
2. *Lack of autocorrelations:* stock returns usually do not exhibit significant autocorrelation, possibly except for very small (intraday) time scales. This property basically implies that there should be no clear structure in the returns dynamics and simple "statistical arbitrage" should not be systemically possible.

3. *Volatility clustering:* volatility measures display positive autocorrelations. This implies that large (small) price fluctuations tend to cluster in time and form volatile (tranquil) market episodes.

To check these properties, we first calculate daily log-returns for randomly chosen simulated price series and for the S&P 500 index. Since it takes about 1,000 trading periods for simulated prices to drift from arbitrary initial values to economically justified levels, we remove these observations from the statistical analysis.

From the visual comparison of simulated and actual returns (see Figure 6) it is clear that the model generates occasional jumps in returns series. It appears that these rare jumps significantly distort properties of simulated returns and hinder meaningful comparison with the actual series. For instance, standard deviation of simulated daily returns is roughly twice as high as that of actual returns (see Table 3 in the Appendix). Daily simulated price fluctuations can reach up to 40 per cent, whereas maximum actual fluctuations were about 10 per cent in the analysed sample. Commensurately, the value at risk (VaR) indicators are significantly higher for the simulated series than the actual series but the difference declines considerably when the VaR threshold is lifted from 1 per cent to 5 per cent. This serves as an indication that the distribution tail behaviour may account for a large part of differences in series properties. Hence, it seems a good idea to dissociate tail behaviour from systemic behaviour of simulated series. For this purpose we arbitrarily exclude observations of returns exceeding 10 per cent in absolute value from the sample as outliers (they constitute about 0.5% of the effective sample).

Figure 6. Simulated and actual daily returns



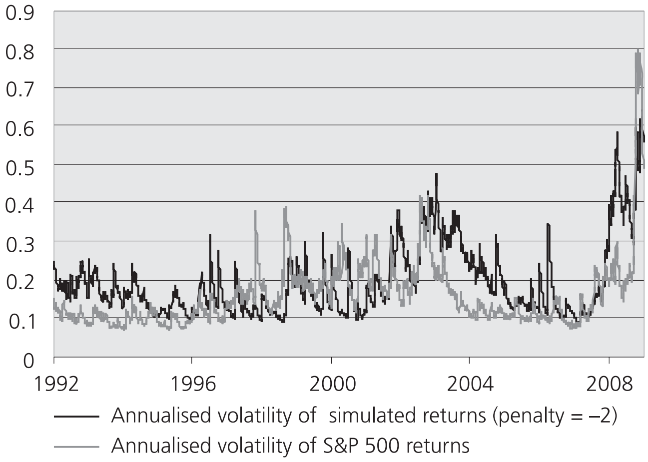
Source: Standard and Poor's data; model simulations.

We turn to the analysis of statistical properties of these resampled series of simulated returns. It turns out that the series exhibit some really nice properties. First of all, simulated returns are non-normally distributed, and distributional characteristics match closely those of the actual returns series. More specifically, simulated returns series with all considered risk parameters definitively fail Jarque-Berra normality tests, along with the actual S&P 500 returns series (see Table 4 in the Appendix). Large kurtosis statistics show that simulated and actual series display heavy tails in line with stylised facts. Also, standard deviations of simulated returns are very similar to the standard deviation of actual S&P 500 returns. One qualitative difference is that in accord with stylised S&P 500 returns distribution displays negative skewness (large negative rewards are more likely), whereas in most cases both original and pruned simulated returns distributions exhibit positive skewness. This is not very surprising bearing in mind both buoyant earnings developments and many accommodating policy shocks in the analysed sample period. Overall, distributional properties of simulated returns are highly realistic.

ASM models often suffer from the deficiency that they generate price series with predictable patterns, which result, e.g., from over- and under-shooting due to agents' herding behaviour. In the current model simulated returns are not significantly autocorrelated. For the reported different simulation runs one-day autocorrelation coefficients vary from 0.04 to 0.08. For comparison, actual S&P 500 returns exhibit negative autocorrelation of 0.08. The small positive autocorrelation of simulated returns may be related to competitive reinforcement learning behaviour, which supports stock price growth following positive returns. Negative autocorrelations with the second lag show that possible over-reactions tend to be corrected immediately. Given the small size of autocorrelation coefficients and the different time scales between actual and simulated series, the correlation of simulated returns is again very much in line with actual returns properties and stylised facts.

Simulated returns' volatility clustering is visible with the naked eye (see Figure 7). This is also confirmed by the statistical analysis. Correlogram tests show that squared returns are indeed positively autocorrelated and correlations die out very gradually as lags increase. For simulated squared returns series one-lag correlations range from 0.12 to 0.19, whereas the autocorrelation for actual squared returns is around 0.34. These correlations suggest that large changes of returns do tend to be followed by further large fluctuations. We also apply a simple GARCH(1,1) model for the simulated returns series and find significant ARCH and GARCH effects in the conditional variance dynamics, which are qualitatively very similar to the actual returns case.

Figure 7. Simulated and actual annualised conditional volatility of stock returns

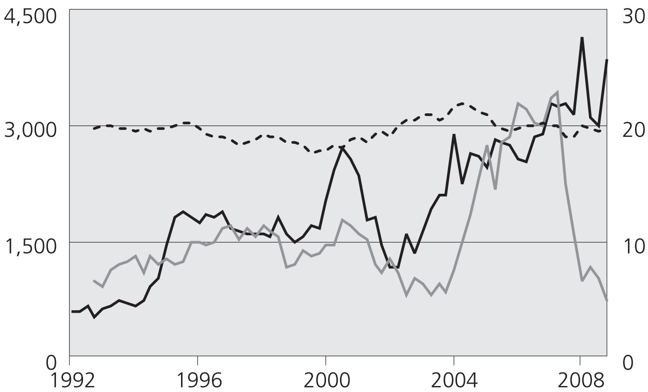


Source: Standard and Poor's data; author's calculations and model simulations.

In general, the model could reproduce the most important stylised properties of actual stock returns very well. This can be to a considerable degree attributed to the fact that the simulated market dynamics strongly depends on simple risk-neutral valuation of fundamentals. And as it turns out, actual stock market movements are tightly related to this simplistic measure of the stock value. Preliminary analysis shows that the model cannot reproduce some the popular stylised facts. For instance, the simulated returns do not display negative correlation with volatility measures and the trading volume is not positively correlated with volatility (see Cont 2001). Further, on the daily time scale simulated returns are not positively correlated with actual returns, though this is not surprising – the correlation becomes significant for quarterly observations of annualised returns.

One remaining important issue is evaluation of our model's results against the background of other calibrated ASM models. The first thing to note is that very few models have attempted to fit model working parameters to actual data in direct estimation procedure (LeBaron 2006), and we did not attempt that here either. Matching emergent properties of artificial stock markets to actual data or to stylised facts also usually gives mixed results. In this context some realistic properties of the current model's simulated market plus qualitative insights about the analysed bubble episodes can be considered as a success. It should be noted that LeBaron (2003) also calibrates his model to U.S. data over quite similar analysis period and his model is also capable of replicating persistent volatility, excess kurtosis and uncorrelatedness of returns. A model developed Farmer and Joshi (2002) in addition to these favourable statistical properties generates "reasonable" long swings away from fundamentals. Iori (2002) and Kirman and Teyssiere (2001), among others, also report uncorrelated returns and persistent volatility of simulated markets.

Figure 8. Simulated market liquidity and window or relevant history compared to market price level



Source: model simulations.

Finally, we touch upon a couple more interesting aspects of the ASM model. Recall our earlier discussion about the possibility that perceived structural breaks of the stock market dynamics can be accompanied by changing sizes of information sets used for earnings forecasts. Model results, quite naturally, give some supporting evidence for these assertions (see Figure 8). For instance, around the time of the first peak of simulated market dynamics agents tend use more recent information but as the stock market returns to its long-term trend agents again rely on longer historical performance. We also note that market activity (quantitative turnover of the stock) quite closely follows simulated price dynamics, though the market liquidity dries out at the end of the sample while the price still remains high.

Conclusion

In our ASM modelling approach we emphasise the importance of the economic content of agent-based models, as overly simplistic or "black-box" modelling do not greatly enhance generative understanding of financial market processes. Though agents' behaviour in the present model is still largely *ad hoc* like in most ASM models, we put a lot of effort in ensuring that their basic behavioural principles accord with commonsensical understanding of sensible investment behaviour in the highly uncertain environment. Agents in this model exhibit inductive behaviour, as they form relatively simple models of the world, act according to them and update those models in order to effectively reach their goals.

Experiments with the model confirm market self-regulation abilities, as different initial market prices tend to converge to similar levels associated with the same underlying fundamentals. The remaining stochastic element of market behaviour suggests that complete market efficiency is hardly possible. The model is also (partly) calibrated to the U.S. financial data, and properties of simulated market dynamics are compared to those of the S&P 500 index. The simulated market exhibits boom-and-bust behaviour and the timing of these booms and busts largely corresponds to actual developments, though simulated structural breaks seem to be more abrupt. A simple empirical analysis and model results also confirm the different nature of the last two global asset bubbles – the technology bubble is more related to unjustified earnings expectations, whereas in the recent asset bubble episode the interest rate environment played a larger role.

Basic statistical properties of simulated market returns closely correspond to known stylised properties of stock returns and actual properties of S&P 500 index returns.

Simulated stock price dynamics displays occasional jumps but, apart from that, simulated returns have a non-normal leptokurtic distribution, they are not significantly autocorrelated and possess the property of volatility clustering. Moreover, many of the descriptive statistics are very similar to those of the actual S&P 500 returns, which is related to the fact that both simulated and actual market dynamics are highly dependent on the dynamics of simple risk-neutral stock fundamentals.

Overall, model results are quite encouraging, though much work is still needed, especially, for developing more transparent, efficient and realistic behavioural methods. Judging the success of a particular model one has to keep in mind that ASM modelling is still in its early stages of development but it does promise to become one of the main alternatives to the crisis-laden neoclassical financial theory.

Table 1

Model parameters

Parameter	Parameter value
General parameters:	
number of agents	200
total number of shares	100
frequency of trading rounds	Daily
frequency of dividend payouts	Quarterly
liquidity ceiling (maximum cash balances allowed)	500
Forecasting dividends:	
dividend forecasting horizon (in quarters)	10
minimum history used for dividend forecasts (in quarters)	6
maximum history used for dividend forecasts (in quarters)	30
probability for choosing data window randomly	0.1
Reinforcement learning:	
learning rate	0.1
exploration rate	0.1
subjective discount parameter of reinforcement learning	0.995
action (incremental adjustment in equation (3))	–0.02; –0.01; –0.005; 0; 0.005; 0.01; 0.02
number of worst-performing agents punished	1
reinforcement signal	Log-returns and penalty

Source: formed by the author.

Table 2

Experimental setting in the case of stationary exogenous processes

Parameter and model setting	Specification
Simulation size:	
length of simulation run (trading rounds *t*)	15,000
number of quarters *q*	241
Dynamics of exogenous variables:	
earnings generating process	$y_q = 100 + 0.05 \cdot 100 \cdot \varepsilon_q$, where $\varepsilon_q \sim N(0,1)$ i.i.d.
dividends	$d_q = 0.4 \cdot y_q$
Term structure of interest rates	$r_{t,t+j} = 0.05 + 0.05 \cdot 0.05 \cdot \varepsilon_t$, where $\varepsilon_t \sim N(0,1)$ i.i.d.
Specific features of experiment runs:	
initial market price	High (50) or Low (5)
penalty for the worst-performing agent	0; –1; –2; –3
loss aversion, %	0; 10; 20; 30

Source: formed by the author.

Table 3

Experimental setting in the case of partial calibration to actual data

Parameter	Parameter value		
Simulation size:			
length of a simulation run (trading rounds *t*)			5,290
number of quarters *q*			84
number of simulation runs in a batch			10
Specific features of experiment runs:	Experiment 1	Experiment 2	Experiment 3
initial market price	Uniformly random in interval (0; 300)		
penalty for the worst-performing agent	–2	–3	–3
loss aversion, %	0	0	10

Source: formed by the author.

Table 4

Properties of simulated and actual returns

	Actual S8P 500	Experiment 1	Experiment 2	Experiment 3
Penalty rate		–2	–3	–3
Loss aversion, %		0	0	10
Initial sample:				
number of observations	4,291	4,291	4,291	4,291
mean	0.000201	0.000417	0.000391	0.000398
median	0.000432	–0.000421	–0.000422	–0.000508
maximum	0.104236	0.295247	0.408859	0.302290
minimum	–0.094695	–0.316977	–0.337488	–0.307772
standard deviation	0.011589	0.020492	0.019549	0.020165
VaR (1%)	0.0313	0.0487	0.0480	0.0465
VaR (5%)	0.0175	0.0204	0.0204	0.0203
skewness	–0.299178	3.373704	3.204439	3.150824
kurtosis	13.41547	76.25842	103.6247	73.59587
Sample with removed extreme observations:				
removed observations	–	24	23	28
remaining observations	4,291	4,267	4,268	4,263
mean	0.000201	–0.000219	–0.000385	–0.000246
median	0.000432	–0.000448	–0.000436	–0.000545
standard deviation	0.011589	0.013961	0.013945	0.013669
skewness	–0.299178	0.167488	0.221399	–0.057599
kurtosis	13.41547	12.47783	13.09445	11.85247
Jarque-Bera	19,459.71	16,080.79	18,253.55	14,013.62
probability	0.000000	0.000000	0.000000	0.000000
one-lag autocorrelation of returns	–0.076	0.044	0.076	0.070
one-lag autocorrelation of squared returns	0.338	0.125	0.188	0.142
ARCH(1)	0.066013	0.043653	0.038483	0.027757
probability	0.000000	0.000000	0.000000	0.000000
GARCH(1)	0.931727	0.951254	0.959871	0.970831
probability	0.000000	0.000000	0.000000	0.000000

Source: formed by the author.

References

Bertsekas D. P., Tsitsiklis J. N. 1996: *A Neuro-Dynamic Programming.* Belmont, MA: Athena Scientific.

Brenner T. 2006: Agent Learning Representation: Advice on Modelling Economic Learning. – *Handbook of Computational Economics* 2. Eds. L. Tesfatsion, K. L. Judd. Amsterdam: Elsevier, 949–1011.

Cont R. 2001: Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. – *Quantitative Finance* 1, 223–236.

Duffy J., Feltovich N. 1999: Does Observation of Others Affect Learning in Strategic Environments? An Experimental Study. – *International Journal of Game Theory* 28, 131–152.

Eshel I., Samuelson L., Shaked A. 1998: Altruists, Egoists and Hooligans in a Local Interaction Model. – *American Economic Review* 88, 159–179.

Farmer J. D., Joshi S. 2002. The Price Dynamics of Common Trading Strategies. – *Journal of Economic Behaviour and Organization* 49, 149–171.

Frankel J. A., Froot K. A. 1987: *Short-Term and Long-Term Expectations of the Yen/Dollar Exchange Rate: Evidence from Survey Data*, NBER Working Paper 0957.

Iori G. 2002: A Microsimulation of Traders Activities in the Stock Market: The Role of Heterogeneity, Agents, Interactions and Trade Frictions. – *Journal of Economic Behaviour and Organization* 49, 269–285.

Jermann U., Quadrini V. 2002: *Stock Market Boom and the Productivity Gains of the 1999s,* NBER Working Paper 9034.

Kahneman D., Tversky A. 1979: Prospect Theory: An Analysis of Decision under Risk. – *Econometrica* 47, 263–291.

Kirchkamp O. 2000: Spacial Evolution of Automata in the Prisoner's Dilemma. – *Journal of Economic Behaviour and Organization* 43, 239–262.

Kirman A. P., Teyssiere G. 2001: Microeconomic Models for Long-Memory in the Volatility of Financial Time Series. – *Studies in Nonlinear Dynamics and Econometrics* 5, 281–302.

LeBaron B. 2003: *Calibrating an Agent-Based Financial Market*. Brandeis University, Working Paper: http://people.brandeis.edu/~blebaron/wps/agcal.pdf

LeBaron B. 2006: Agent-Based Computational Finance. – *Handbook of Computational Economics* 2. Ed. L. Tesfatsion, K. L. Judd. Amsterdam: Elsevier, 1189–1233.

Martinez-Jaramilo S. 2007: *Artificial Financial Markets: An Agent Based Approach to Reproduce Stylized Facts and to Study the Red Queen Effect.* PhD Thesis, University of Essex.

Mitchell T. 1997: *Machine Learning.* New York: McGraw Hill.

Ramanauskas T. 2008: Agent-Based Financial Modelling as an Alternative to the Standard Representative-Agent Paradigm. – *Pinigų studijos* 2, 5–21.

Ramanauskas T., Rutkauskas A. V. 2009: *Building an Artificial Stock Market Populated by Reinforcement-Learning Agents.* Lietuvos bankas, Working Papers Series, No 6.

Sutton R. S., Barto A. G. 1998: *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

Watkins C. J. 1989: *Learning from Delayed Rewards.* PhD Thesis, Cambridge University.

Gauta 2009 m. sausio mėn.
Priimta spaudai 2009 m. balandžio mėn.

Santrauka

DIRBTINĖS AKCIJŲ RINKOS MODELIO EMPIRINĖ VERSIJA

Tomas Ramanauskas

Šiame straipsnyje pristatomas dirbtinės akcijų rinkos modelis, pagrįstas heterogeniškų jos dalyvių kompleksine sąveika. Modelis iš dalies kalibruojamas, naudojant empirinius duomenis. Juo siekiama analizuoti rinkos nepusiausvyros kitimą, tirti rinkos savireguliacijos galimybes bei sistemines savybes ir, pasitelkiant JAV finansų rinkos pavyzdį, gerinti pastarųjų dešimtmečių finansinių burbulų susidarymo ir sprogimo aplinkybių generatyvinį aiškinimą.

Sudarant modelį, didelė reikšmė teikiama ekonominiu požiūriu priimtinai individualių rinkos dalyvių, arba veikėjų *(agents)*, elgsenai: jų elgsena induktyvi, veikėjai kuria paprastus aplinkos modelius, jais vadovaudamiesi priima investicinius sprendimus ir, atsižvelgdami į tai, kaip jiems sekasi siekti savo strateginių ekonominių tikslų (maksimalios investicinės grąžos), tuos sprendimus persvarsto ir tikslina. Rinkos dalyvių individualus prisitaikymas

pagrįstas stiprinamuoju mokymusi, tiksliau – dirbtinio intelekto teorijos pateikiamu Q mokymosi algoritmu. Dirbtinio intelekto teorijos sritis – stiprinamasis mokymasis ir jo algoritmai ekonominio pobūdžio problemoms spręsti naudojami retai, tačiau tai yra gana perspektyvi ekonomikos veikėjų elgsenos modeliavimo priemonė.

Dar vienas modelio ypatumas – gana detalus rinkos mechanizmo ir prekybos proceso aprašymas. Eksperimentai su modeliu patvirtina tam tikras dirbtinės rinkos savireguliacijos galimybes. Naudojant JAV ekonominius duomenis kalibruojamo modelio imitacinė rinka pasižymi tam tikrais finansinių burbulų susidarymo ir sprogimo aspektais, kurie iš esmės atitinka faktinius. Modeliavimo rezultatų palyginimas su JAV akcijų rinkos faktine raida sudaro prielaidas įžvalgoms, kad XXI a. pirmojo dešimtmečio pradžios technologijos burbulą iš esmės lėmė nepagrįsti įmonių pelno lūkesčiai, o pastarųjų metų finansinis burbulas labiau sietinas su ypač palankia palūkanų normų aplinka.