VILNIUS GEDIMINAS TECHNICAL UNIVERSITY

Tomas RAMANAUSKAS

STOCK MARKET SIMULATION BASED ON INTERACTION OF HETEROGENEOUS LEARNING AGENTS

DOCTORAL DISSERTATION

SOCIAL SCIENCES,
ECONOMICS (04S)

VGTU
LEIDYKLA
Vilnius TECHNIKA 2010

Doctoral dissertation was prepared at Vilnius Gediminas Technical University in 2005–2010.

Scientific Supervisor

Prof Dr Habil Aleksandras Vytautas RUTKAUSKAS (Vilnius Gediminas Technical University, Social sciences, Economics – 04S).

http://leidykla.vgtu.lt

VGTU leidyklos TECHNIKA 1758-M mokslo literatūros knyga

ISBN 978-9955-28-574-8

© VGTU leidykla TECHNIKA, 2010
© Tomas Ramanauskas, 2010
tomas_ramanauskas@yahoo.com

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

Tomas RAMANAUSKAS

HETEROGENINIŲ BESIMOKANČIŲ AGENTŲ SĄVEIKA PAGRĮSTAS IMITACINIS AKCIJŲ RINKOS MODELIAVIMAS

DAKTARO DISERTACIJA

SOCIALINIAI MOKSLAI,
EKONOMIKA (04S)

VGTU
LEIDYKLA
Vilnius TECHNIKA 2010

Disertacija rengta 2005–2010 metais Vilniaus Gedimino technikos universitete.

Mokslinis vadovas

prof. habil. dr. Aleksandras Vytautas RUTKAUSKAS (Vilniaus Gedimino technikos universitetas, socialiniai mokslai, ekonomika – 04S).

Abstract

The main object of this study is the stock market, seen as a complex system constituted of basic elements (securities, trading infrastructure and atomistic heterogeneous investors) and process flows (forecasting, investment decision making, trade execution, maintenance of financial records, etc.). More specifically, we explore the possibilities to use specific simulation models as tools for generating economically interesting and adequate systemic behaviour, analysing properties of the actual stock market and explaining certain market phenomena. We propose two specific artificial stock market (ASM) models, based on inductive, competitive reinforcement-learning behaviour of atomistic heterogeneous agents in highly uncertain environment. These models are used as a structured framework for the analysis of market self-regulation abilities, efficiency, determinants of price formation and investors' attitude to risk, emergent properties of the financial market and possible determinants of asset price bubbles. One model is calibrated to actual financial data.

One of distinctive features of the proposed modelling approach is a strong emphasis on economic behaviour of individual agents. In the proposed models boundedly rational agents base their decisions on economic considerations, such as estimation of discounted earnings and comparison of returns on different investment strategies, and pursue forward-looking behaviour in highly uncertain environment. Agents' individual adaptation, intertemporal decision making and forward-looking behaviour in the multi-agent setting is governed by reinforcement learning technique borrowed from the field of machine learning. To our knowledge, this work is one of the first attempts to apply the reinforcement learning techniques in an ASM model. Also, this is one of the first full-fledged artificial stock market models in the Lithuanian scientific literature.

The Dissertation is comprised of an introductory chapter, five main chapters, a concluding chapter, two appendices describing model setups and results, and two appendices with program codes for the proposed models. Chapters 1-3 contain theoretical, methodological and discussion material. In Chapter 1 we give a general discussion about agent-based financial models as an alternative and a complement to standard financial theories. A brief ASM literature review and presentation of some specific ASM design issues are provided in Chapter 2. Chapter 3 is devoted to intuitive presentation of basic principles of the standard reinforcement learning and contains discussion on its relevance in the context of financial modelling. In Chapter 4 we spell out the general analytical framework of the simulation model and conduct some simulation exercises. Chapter 5 contains a parsimonious, suitable for empirical calibration version of the model and simulation results.

Reziumė

Šio tyrimo objektas yra akcijų rinka, kuri suvokiama kaip kompleksinė sistema, sudaryta iš bazinių elementų (vertybinių popierių, prekybos infrastruktūros ir atomistinių heterogeninių investuotojų) ir procesų (prognozavimo, investicinių sprendimų priėmimo, finansinių sąskaitų vedimo ir t. t.). Konkrečiau tariant, šiame darbe nagrinėjamos galimybės panaudoti konkrečius imitacinius modelius siekiant generuoti ekonominiu požiūriu įdomią sistemos dinamiką, tirti faktinės rinkos savybes bei paaiškinti kai kuriuos rinkos reiškinius. Šiuo tikslu pasiūlyti du dirbtinės akcijų rinkos modeliai, pagrįsti heterogeninių agentų induktyvia, konkurencine skatinamojo mokymosi elgsena dideliu neapibrėžtumu pasižyminčioje aplinkoje. Šie modeliai sudaro struktūrizuotą analitinį pagrindą rinkos savireguliacijos galimybėms, efektyvumui, kainos formavimosi veiksniams, investuotojų požiūriui į riziką, kylančioms rinkos savybėms bei galimiems finansinių burbulų veiksniams tirti. Vienas iš šių modelių yra iš dalies pritaikytas empiriniams finansiniams duomenims.

Išskirtinis pasiūlytų dirbtinės akcijų rinkos modelių ypatumas yra tas, kad teikiama didelė svarba ekonominiu požiūriu pagrįstam agentų elgsenos modeliavimui. Šiuose modeliuose riboto racionalumo agentai savo sprendimus grindžia ekonomine logika, pavyzdžiui, diskontuotų pajamų srautų vertinimu ar alternatyvių investicinių strategijų lyginimu, o jų elgsena neapibrėžtumu pasižyminčioje aplinkoje yra orientuota į ateities galimybių vertinimą. Individuali agentų adaptacija bei tarplaikiniai investiciniai sprendimai pagrįsti skatinamojo mokymosi algoritmu, perimtu iš mašinų mokymosi literatūros. Šis tyrimas yra vienas iš pirmųjų bandymų pritaikyti skatinamojo mokymosi algoritmus imitaciniame akcijų rinkos modelyje. Tai yra vienas iš pirmųjų pilnos apimties imitacinių akcijų rinkos modelių Lietuvos mokslinėje literatūroje.

Disertaciją sudaro įvadinis skyrius, penki pagrindiniai skyriai, išvadų ir rezultatų skyrius, du priedai aprašantys modelio nustatymus bei eksperimentų rezultatus ir du priedai, kuriuose išdėstytas modelių programinis kodas. Pirmame skyriuje pateikiama bendra diskusija apie agentais pagrįstą finansinį modeliavimą kaip alternatyvą standartinėms neoklasikinėms teorijoms. Atranki literatūros apžvalga bei svarbūs imitacinės akcijų rinkos modelių kūrimo aspektai aptariami antrame skyriuje. Trečiame skyriuje pristatomi baziniai skatinamojo mokymosi algoritmo (tiksliau, vadinamojo Q-mokymosi algoritmo) principai bei aptariamos galimybės jį taikyti vykdant finansinį modeliavimą. Ketvirtame skyriuje pateikiamas bendras imitacinio modelio pagrindas bei atliekami kai kurie imitaciniai eksperimentai. Penktame skyriuje pristatoma glaustesnė, empiriniam taikymui tinkama modelio versija bei aprašomi imitacinių eksperimentų rezultatai.

Notation

Symbols

$a_{i,t}^{p}$ – agent *i*'s individual price adjustment factor;

$a_{i,t}$ – agent *i*'s individual price adjustment factor in the parsimonious model;

a_t – action taken at time *t*;

$a_{i,y}^{div}$ – agent *i*'s individual dividend expectation adjustment factor;

$b_{i,t}$ – indicator variable showing whether agent *i* is willing to buy or sell the stock;

c – prespecified fractional trading cost;

d_y – annual dividend payout;

$d_{i,y+n}^{e}$ – agent *i*'s individual *n*-year-ahead dividend forecast;

d_t – dividend pay-out during specific trading session *t*;

$h_{i,t}^{0}$ – agent *i*'s stock holdings (number of shares) just before trading session *t*;

$h_{i,t}^{1}$ – agent *i*'s stock holdings (number of shares) immediately after trading session *t*;

y – time period (year);

$y_{i,q+j}^{e}$ – individual *j*-quarter-ahead forecast of quarterly corporate earnings;

y_q – corporate earnings in quarter *q*;

$m_{i,t}^{0}$ – agent *i*'s cash balances just before trading session *t*;

$m_{i,t}^{1}$ – agent *i*'s cash balances immediately after trading session *t*;

$p_{i,t}^{q}$ – stock price quoted by individual *i* in trading session *t*;

$P_{ss'}^{a}$ – probability of transition to state s' from current state *s* given action *a*;

p_t – average transaction price (market price) in trading session *t*;

$Q^{\pi}(s,a)$ – action-value function for the state-action pair under policy π;

q – time period (quarter);

$\bar{r}^{monthly}$ – a predetermined one-period return on bank account;

$\bar{r}$ – constant interest rate;

$r_{i,y+n}^{d}$ – reinforcement signal to agent *i* associated with n-year-ahead dividend forecast;

$r_{i,t+1}$ – reinforcement signal to agent *i* in the refined model;

$r_{i,t+1}^{p}$ – reinforcement signal to agent *i* associated with stock value assessment;

$r_{i,t}^{return}$ – individual return on wealth following trading session *t*;

$r_{q,q+j}^{f}$ – market yield on the government bond maturing in j quarters;

$R_{ss'}^{a}$ – expected reward signal given transition from state *s* to s' and action *a*;

r_t – reinforcement signal at time *t*;

s_t – environment state at time *t*;

t – time period;

$V^{\pi}(s)$ – state-value function for action policy π in state *s*;

$v_{i,t}^{reserve}$ – individual reservation price at time *t* (trading session *t*);

v_t – *t*-th iterative approximation of action-value function;

$v_{i,y}^{fund}$ – agent *i*'s perceived risk-neutral present value of expected dividend flows;

$w_{i,j,t}^{1}$ – expected end-of-period wealth of agent *i* for *j*-th feasible price quote;

$x_{j,t}$ – *j*-th feasible price quote in trading session *t*;

α – the learning rate parameter in the reinforcement learning algorithm;

$\beta_{\bullet,i}^{SR}$ – individual coefficients in the short-run regression of corporate earnings;

$\beta_{1,i}^{LR}$ – individual intercept coefficient in the long-run regression of corporate earnings;

$\beta_{2,i}^{LR}$ – individual slope coefficient in the long-run regression of corporate earnings;

γ – discounting parameter in the reinforcement learning algorithm;

ε – probability of choosing non-greedy action;

$\varepsilon_{i,q}^{LR}$ – ordinary least squares residual in the long-run regression of corporate earnings;

$\vec{\theta}$ – column vector of parameter matrix Θ associated with action *a*;

λ_1 – smoothing factor in dividend expectations;

λ_2 – smoothing factor in perceived present value calculation;

$\pi(s, a)$ – probability of taking action *a* in state *s* (action policy);

$\vec{\phi}_s$ – vector of state features;εθικλμνξπστφ

Θ – parameter matrix in the action-value function approximation.

Abbreviations

EMH – efficient market hypothesis;
ABF – agent-based finance;
ASM – artificial stock market;
EWMA – exponentially weighted moving average;
VaR – value at risk.

Table of Contents

INTRODUCTION 1
Research motivation 1
Research problem 2
The object of research 3
The goal of the Dissertation 3
Research tasks 3
Research methodology 4
Scientific novelty of research 4
Practical importance of research 5
Defended theses of the Dissertation 5
Structure of the Dissertation 6
Approval of research results 6

1. AGENT-BASED FINANCIAL MODELS VERSUS THE REPRESENTATIVE-AGENT PARADIGM 7
1.1. General discussion of homogeneity, perfect rationality and market efficiency 8
1.2. The new paradigm – markets as complex agent-based systems 11
1.3. Concluding summary of Chapter 1 14

2. ARTIFICIAL STOCK MARKET LITERATURE REVIEW AND DISCUSSION OF MODEL DESIGN ISSUES 17
2.1. ASM building principles and problems 17
2.2. Brief ASM literature review 21

2.2.1. ASM models based on random, heuristic and hard-wired behavioural rules 22
2.3.1. ASM models based on intelligent adaptation 23
2.4. Related research in Lithuanian scientific literature 27
2.5. Concluding summary of Chapter 2 28

3. REINFORCEMENT LEARNING IN THE CONTEXT OF ARTIFICIAL STOCK MARKET MODELLING 31
3.1. Economic rationale for choosing reinforcement learning method 32
3.2. Basic principles of reinforcement learning 33
3.3. Potential for application of reinforcement learning technique in economics and finance 40
3.4. Concluding summary of Chapter 3 41

4. BUILDING ARTIFICIAL STOCK MARKET POPULATED BY REINFORCEMENT-LEARNING AGENTS 43
4.1. Description of the ASM model 44
4.1.1. General market setting and model's main building blocks 45
4.1.2. Forecasting dividends 47
4.1.3. Estimating fundamental the stock value and reservation prices 48
4.1.4. Making individual trading decisions 49
4.1.5. Learning and systemic adaptation in the model 54
4.2. Simulation results 57
4.3. Concluding summary of Chapter 4 61

5. ARTIFICIAL STOCK MARKET MODEL VERSION SUITABLE FOR EMPIRICAL ANALYSIS 63
5.1. Description of the model 64
5.1.1. Model's main building blocks 65
5.1.2. General market setting 66
5.1.3. Forming individual forecasts of corporate earnings 67
5.1.4. Estimating fundamentals and reservation prices 70
5.1.5. Agents' reinforcement learning 72
5.1.6. Carrying out trades via the centralised exchange 74
5.2. Simulation in stationary environment and examination of market self-regulation ability 75
5.3. Applying the model to the actual data 78
5.3.1. Model data and background discussion 79
5.3.2. Comparison of simulated and actual market dynamics 83
5.3.3. Analysis of simulated returns properties 85
5.4. Concluding summary of Chapter 5 89

GENERAL CONCLUSIONS 91
Concluding summary and comments 91
Main conclusions 94

REFERENCES 97

AUTHOR'S PUBLICATIONS RELATED TO THE DISSERTATION RESEARCH 105

APPENDICES 107
Appendix A. Parameter setting and experimental results of the initial model 108
Appendix B. Parameter setting and experimental results of parsimonious version of the model 116
Appendix C. Program code for the initial version of the model.....119
Appendix D. Program code for the parsimonious version of the model 137

List of Figures

Fig. 2.1. Basic process flow in ASM models 19
Fig. 3.1. Basic Q-learning algorithm 38
Fig. 3.2. Gradient-descent approximation Q-learning algorithm 39
Fig. 4.1. Typical estimated demand and supply schedules in an upward-moving market 52
Fig. 4.2. Typical dynamics of percentage stock returns and percentage changes in liquidity in a constant dividend case 60
Fig. 5.1. Simulated market price dynamics from different initial price levels (stationary external environment; penalty rate = –1; no loss aversion) 76
Fig. 5.2. Simulated market dynamics for different loss aversion parameters (stationary external environment; penalty rate = – 2) 78
Fig. 5.3. Dynamics of S&P 500 index and official timing of U.S. recessions 79
Fig. 5.4. Quarterly corporate earnings of companies included in S&P 500 index (bn U.S. dollars) 81
Fig. 5.5. Corporate earnings as a fraction of market value compared to riskfree returns 82
Fig. 5.6. Average simulated estimate of fund. value compared to S&P500 index 83
Fig. 5.7. Simulated market prices compared to S&P 500 index 84
Fig 5.8. Simulated and actual daily returns 86
Fig. 5.9. Simulated and actual annualised conditional volatility of stock returns 87
Fig. 5.10. Simulated market liquidity (no. of transactions) and window or relevant history (in quarters) compared to market price level 89
Fig. A1. Selected graphs of Experiment 1 111
Fig. A2. Selected graphs of Experiment 2 112

Fig. A3. Selected graphs of Experiment 3 113
Fig. A4. Selected graphs of Experiment 4 115

List of Tables

Table 2.1. Comparison of the two broad categories of ASM models 24
Table 4.1. Main building blocks of the ASM model..... 46
Table 5.1. Main building blocks of the empirical ASM model..... 66
Table A1. Key parameter settings of the ASM model 108
Table A2. Specification of model experiment runs..... 109
Table A3. Basic descriptive statistics of simulation experiments 109
Table B1. Model parameters..... 116
Table B2. Experimental set-up with stationary exogenous processes..... 117
Table B3. Experimental setting in the case of calibration to actual data..... 117
Table B4. Properties of simulated and actual returns..... 118

Introduction

Research motivation

The world economy is in the midst of a global financial crisis, caused by a mix of a global asset price bubbles, overwhelming irrational exuberance and systemic mistakes of economic agents and financial market participants. Against this background, the neoclassical financial theory – based on strong assumptions about the market efficiency, rational representative agent and market clearing achieved through Walrasian tâtonnement processes – seems to be losing touch with reality. To understand the increasingly anomalous market behaviour it might not suffice introducing market imperfections or changing other aspects of standard models. There is mounting evidence that the paradigm shift might be necessary.

Unfortunately, there are no satisfactory alternatives yet but, with growing computing power, modelling possibilities expand and new promising frontiers of research emerge. One of them could be agent-based computational finance. Agent-based financial models take account of some fundamental features inherent for the real world financial markets, such as systemic complexity, agent heterogeneity, bounded rationality and complex interaction of agents. More generally, these simulation models give researchers great flexibility to model interesting features of real world phenomena. This will eventually allow very

realistic modelling of financial markets, goods markets or the economy as a whole. Before this vision materialises, however, a major breakthrough in realistic modelling of intelligent, but boundedly rational human behaviour is needed. It can only be achieved by blending and expanding advancements of the economic theory, cognitive psychology and artificial intelligence theory (LeBaron, 2003).

Research problem

One of the major tasks of the financial research in general is to enhance our generative understanding of the actual functioning of financial markets. In the broadest sense, this work also aims at generative, bottom-up understanding of what principal elements and processes constitute stock markets and how complex market behaviour may emerge. Within the scope of artificial stock market modelling, the research problem can be articulated through a number of generally unanswered questions: what are the crucial elements and processes of financial markets, what are the main drivers behind investors' decisions, what are emergent properties of the simulated market and how they conform with actual market properties and stylised empirical facts, what can be said about the formation and stability of market equilibrium, how and why financial bubbles form, etc.

We hold the view that whilst the relaxation of strong neoclassical assumptions about the principles of market functioning is a big step forward in the quest of answering the above questions, researchers still face a huge challenge of making the agents' behaviour in these models more "economic". In other words, agents should exhibit economically interesting behaviour and retain elements of economic reasoning rather than constitute mere collections of behavioural rules. Ensuring that agents' behaviour has sound economic rationale and is validated by experimental evidence would greatly increase the credibility of model results but this largely remains an unresolved problem (Duffy, 2006).

Another caveat is that systemic adaptation in most agent-based financial models relies on evolutionary search algorithms. This means that systemic dynamics, e. g. trading and market price formation is generated by simply ensuring the sufficient variety of investment strategies and inducing some sort of evolutionary selection of strategies in favour to those that give highest utility to individuals. Such approach often downplays the importance of individual behaviour, which is often assumed to be driven by simplistic rules. Also, these algorithms generally do not support forward-looking behaviour except special cases, in which agents try to achieve myopic one-period optimisation. Unlike neoclassical financial theories, most existing agent-based models are not well-

suited to model the intertemporal choice and hence miss a crucial aspect of financial decision making.

The object of research

The main object of this study is the stock market, seen as a complex system constituted of basic elements (securities, trading infrastructure and atomistic heterogeneous investors) and process flows (forecasting, investment decision making, trade execution, maintenance of financial records, etc.). More specifically, we explore the possibilities to use specific simulation models as tools to generate economically interesting and adequate systemic behaviour, analyse the properties of the actual stock market and explain certain market phenomena (e. g. bubble formation).

The goal of the Dissertation

The goal of the Dissertation is to:

- Propose specific agent-based artificial stock market (ASM) models, based on inductive, competitive reinforcement-learning behaviour of atomistic heterogeneous agents in highly uncertain environment;
- Use the models as a structured framework for the analysis of market self-regulation abilities, efficiency, determinants of price formation and investors' attitude to risk, emergent properties of the financial market and possible determinants of asset price bubbles.

Research tasks

In this work we set up the following research tasks:

1. Present main concepts and principles of agent-based modelling, discuss the main differences between the standard neoclassical financial theories and the agent-based approach and provide an overview of the literature on the artificial stock market modelling.
2. Present the main principles of the reinforcement learning methodology and discuss issues related to its implementation in the context of financial modelling.

3. Develop a conceptual framework for the agent-based analysis of the artificial stock market, discuss behavioural principles and market processes in the simulation model.
4. Develop a parsimonious version of the agent-based model of the stock market populated by heterogeneous reinforcement-learning agents, conduct simulation experiments, analyse emergent properties of the simulated market and determine possible drivers behind financial bubble formation.
5. Calibrate the model to empirical financial data and compare simulated market's properties to those of the actual stock market and to stylised empirical properties.

Research methodology

This research relies on simulation-based scientific enquiry, descriptive statistics, econometric modelling, theories of economic decision making, and the reinforcement learning algorithm (Q-learning) taken from the machine learning research.

Scientific novelty of research

The proposed approach to the ASM modelling is quite independent and has a number of novel features. We strive to overcome some of the abovementioned problems of agent-based modelling without sacrificing their merits, and this requires seeking a good balance between systemic adaptation (through competition) and individual adaptive behaviour. For this purpose we put a much stronger emphasis on economic behaviour of individual agents than in most agent-based models. In our research, boundedly rational agents base their decisions on economic considerations, such as estimation of discounted earnings and comparison of returns on different investment strategies, and pursue forward-looking behaviour in highly uncertain environment. They also take account of the uncertain impact of other market participants' actions on the market price.

Individual adaptation, intertemporal decision making and forward-looking (non-myopic) behaviour in the multi-agent setting is governed by reinforcement learning technique borrowed from the field of machine learning. To our knowledge, our work is one of the first attempts to apply the reinforcement learning techniques in an ASM model. The proposed models apparently are the

first attempts in Lithuanian scientific literature to develop a full-fledged artificial stock market.

There are some other interesting and novel features in the proposed models. Agents' interaction in our models is highly autonomous, as we do not make any strong assumptions about aggregate supply, demand or price adjustments. The proposed framework enables us to formulate and examine ideas about possible determinants of bubble formation. Overall, this research contributes to the agent-based modelling by offering some interesting modelling ideas and helping to narrow the gap between agent-based models of financial markets and more conventional approaches to financial modelling.

Practical importance of research

Though theoretical in essence, the proposed models have some practical value in that enable investors and policy makers to gain qualitative insights about bubble formation mechanisms and examine the effects of interesting shocks to the stylised simulated financial market. Reinforcement learning techniques and some other ideas from the current research can be extended for application in different financial research areas, such as the analysis of portfolio management, stability of financial systems, bank runs and financial panic, etc.

Defended statements of the Dissertation

1. The neoclassical financial theory is based on restrictive assumptions and focuses on equilibrium relationships, and thereby it fails to provide generative explanation of financial market functioning. In contrast, agent-based financial modelling is better suited for this purpose as it recognises that systemic complexity, investor heterogeneity, bounded rationality and inductive behaviour are salient features of financial markets.
2. The proposed basic ASM model can serve as a plausible framework for generative analysis of financial market processes, such as price determination, making of trading decisions, expectations formation, etc.
3. The proposed calibrated model version is capable of generating the artificial stock market with simulated returns properties closely matching actual financial data and stylised properties of stock market returns.

4. The simulated market in the calibrated model version has a tendency of self-organisation, i. e. transition of the complex system to the state of equilibrium price; individual adaptation and evolutionary selection are crucial for this sort of systemic behaviour in the model.
5. The proposed models can be used to demonstrate that plausible financial bubble formation mechanisms include excess liquidity, contagious earnings forecast errors and competitive pressures to reap short-term benefits from foreseeable price trends during bubble formation episodes.
6. Reinforcement learning is conceptually suitable for modelling agents' inductive behaviour in an artificial stock market model, even though it cannot guarantee optimal behavioural strategies.

Structure of the Dissertation

The Dissertation is comprised of introduction, five main chapters, chapter for general conclusions and discussion of results, list of references, list of author's publications and four appendices. In Chapter 1 we give a general discussion about agent-based financial models as an alternative and a complement to standard financial theories. A brief ASM literature review and presentation of some specific ASM design issues are provided in Chapter 2. Chapter 3 is devoted to intuitive presentation of basic principles of the standard reinforcement learning and contains discussion on its relevance in the context of financial modelling. In Chapter 4 we spell out the general analytical framework of the simulation model and conduct some simulation exercises. Chapter 5 contains a parsimonious, suitable for empirical calibration version of the model and simulation results.

Approval of research results

Most of the material of this dissertation is very much interlinked with three scientific papers prepared by the author (in collaboration with Dissertation Supervisor Professor Aleksandras Vytautas Rutkauskas) during the period of his PhD. studies, namely Ramanauskas (2008), Rutkauskas and Ramanauskas (2009) and Ramanauskas (2009). The author has given several presentations on topics related to dissertation research at an international conference (Ramanauskas, 2007), in workshops held at Stockholm School of Economics in Riga, at Vilnius Management Academy and at student seminars.

1

Agent-based financial models versus the representative-agent paradigm

Neoclassical capital market theories regard financial markets as extremely efficient price determination mechanisms. Moreover, trading processes themselves are effectively excluded from the analysis by applying strong market clearing assumptions. Now one can observe the gradual paradigm shift in the financial literature, as financial markets are being increasingly viewed as complex dynamical systems, consisting of interacting atomistic agents whose complex interaction and individual learning result in some systemic adaptation but not necessarily high market-level efficiency. The ongoing paradigm shift is by no means a merely academic debate. It is of great importance for market participants and policy makers. It is pretty obvious that in recent years economists' blind belief in nearly perfect markets' self-regulation abilities outshined the premonition of the looming global financial catastrophe and arguably even paved the way for it. Hence, the current crisis offers economic researchers a good reason to devote more effort for enhancing generative understanding of market processes rather than concentrate on equilibrium relations.

Much of the mainstream financial theory builds on the efficient market hypothesis (EMH) and the rational representative agent paradigm. These presumptions have clearly played a crucial role in shaping the widely accepted

understanding of risk, determinants of asset prices, portfolio management principles, etc. Yet there is a growing need to recognise that the gap between this idealisation and reality may be too substantial for the theory to grasp correctly the essence of the functioning of financial markets, as the standard theory arguably abstracts from the salient features of examined phenomena. The central question is whether the financial market can be seen as a perfect (or nearly perfect) price determination mechanism. There are a lot of theoretical caveats and empirical anomalies associated with standard financial theories, strong assumptions of perfect rationality and the efficient market hypothesis.

In this chapter we discuss the problem issues of standard financial market modelling and advocate the position that some of them can eventually be solved by taking alternative, agent-based modelling approach. We present main concepts and principles of agent-based modelling and highlight the main differences between the standard neoclassical financial theories and the agent-based approach.

1.1. General discussion of homogeneity, perfect rationality and market efficiency

Theoretically, the homogeneous agent assumption is far from innocuous, as at the heart of finance lies the collective discovery of securities prices in the process of trading, i. e. as a result of the interaction of heterogeneous agents. It is a plain fact, which hardly requires any scientific inquiry, that investors have a bewildering variety of views, expectations and preferences. They possess different amounts of information, and their financial decisions vary greatly in the level of sophistication. In reality, the financial market awakes to life and trading takes place exactly owing to this heterogeneity of market participants. An assumption that each individual, or the aggregate market behaviour, can be approximated by some average or a fictitious representative agent inevitably leads to the loss of a large degree of freedom. By assuming this, one clearly risks attributing effects of changes in individual agents' perceptions and strategies to something like the representative agent's consumption smoothing preferences.

In reality it is much, much more complex. For example, the standard line of thinking suggests that individuals invest in risky assets so as to optimise their consumption patterns. The big question is whether they are able to do that effectively if the environment is highly uncertain. Assuming that they can, they will still have widely differing valuations of expected payoffs, owing to their different preferences and risk tolerance. Moreover, huge amount of idiosyncratic changes in investors' beliefs, preferences, needs, etc., force investors to adjust their investment positions. As a result, aggregate supply and demand shift, and

market prices of risky assets change commensurately. The magnitude of this change is largely unpredictable because in reality there is no way of knowing idiosyncratic factors affecting each investor's asset supply and demand curves. In the short run, stock prices are arguably more affected by these idiosyncratic shocks than by relatively infrequent news on the structural changes of the processes materially affecting fundamentals – this idea can be traced back to Keynes (1936)[1]. In this light market efficiency, as well as congruence between fundamentals and market valuation become very vague concepts: what are those fundamentals if everyone can have different and not necessarily erroneous beliefs about what an asset is worth *to them*? It is immensely doubtful that is is valid to claim that the "true" fundamental value of an asset is simply some average of all individual valuations. Such concept can only be a merely theoretical construct, and recent global financial developments have forcefully demonstrated that the average market valuation may diverge dramatically from any sort of fundamentals.

Even stronger is the perfect rationality assumption. It is obvious that, as already noted by Simon (1957) half a century ago, individuals act in a highly uncertain environment and their natural computing abilities are limited, while information search and analysis are costly and time consuming. All of this implies that even if perfect rationality was feasible in the information collection, processing and decision making sense, it would simply be too costly economically. Paradoxically, it is hardly rational to attempt being perfectly rational. Moreover, a large thread of literature of psychology and behavioural finance instigated by laboratory experiments of Kahneman and Tversky (1973) and Tversky and Kahneman (1974) suggest that economic behaviour is often better explained by simple heuristic rules and irrational biases rather than by dynamic optimisation.

The perfect rationality assumption is particularly strong and restrictive if applied in the context of financial market modelling. This is related to an inherently large impact of expectations on the aggregate financial market behaviour. Recall in this context Keynes's famous "beauty contest" idea that it often pays to prefer crowd wisdom to own beliefs about fundamentals, and as we argued above there are no market forces that could ensure that the crowd is always unerring. Moreover, every market participant has some marginal impact on these price fluctuations and their expectations about likely price changes may become self-fulfilling. Any signal, such as good news related to a specific stock, may lead to investors' coincident actions. This often triggers changes in the stock price in a predictable direction, which implies that immediately following the news it can become optimal for short-term speculators to buy the stock

[1] See also Cutler, *et al.* (1989).

irrespective of actual fundamentals. There is nothing to prevent under- or over-reactions to the news, so once again it is highly unrealistic to assume that the actual market price always coincides with some fundamental value. In general, self-fulfilling expectations may lead to multiple sunspot equilibria, which are not consistent with rational expectations by definition. In other words, Muth's (1961) rational expectations hypothesis can simply be seen as an elegant way to exclude *ad hoc* forecasting rules and market psychology from economic modelling (Hommes, 2006) but due to the self-referential nature of predictions they may be deductively indeterminate. In reality market participants are more likely to form expectations inductively (Arthur, 1995) – subjective expectations are formed, tested and changed dynamically, as market conditions change and market participants gain experience or interpret (possibly erroneously) plentiful information signals.

We all know from our own experience that there is no conceivable mechanism ensuring that people's economic or social behaviour is perfectly rational. Yet proponents of the EMH hypothesis argue that perfect rationality can be an emergent feature of the financial market[2]. There are claims, for instance, that the existence of arbitrage traders, evolutionary competition and generally offsetting each other noise traders' bets may ensure that securities prices always reflect fundamentals correctly. The idea (which is also known in the literature as the Friedman hypothesis) that poor performance drives non-rational investors out of the market is indeed appealing. However, investment in stocks and some other securities is not a zero-sum game. Stock prices generate positive returns in the long run in the overwhelming majority of cases. Hence, it is not clear why non-rational investors, especially passive investors, should "die out" – they may well enjoy decent returns to their less-than-rational (e. g. passive) investment strategies. Moreover, their army is constantly replenished with new inexperienced and hence non-rational investors. The argument of a negligible impact of emotional and non-rational traders can also be challenged. It is exactly them, rather than "fundamentalist" traders, who are more likely to react to non-fundamental headline news and push market prices in the predictable direction, acting as a powerful market-moving force and thereby imposing "rules of the game". Moreover, it is well known that sophisticated traders, instead of acting as a stabilising force, may try to exploit the resultant predictable market movements. For instance, Frankel and Froot (1986) conclude from their survey that investors often recognise a considerable price deviation from their perceived fundamentals but nevertheless they find it logical to follow the trend until it reaches some turning point.

[2] Emergent features are systemic features that cannot be deduced by simply scaling individual behaviour; see Chen and Yeh (2002b) for discussion.

There are also a number of empirical problems with traditional financial models based on perfect rationality and EMH assumptions. Financial literature discusses quite a few empirical anomalies, i. e. empirical regularities that are not explained by the theory. Probably the most famous one is the equity premium puzzle raised by Mehra and Prescott (1985). Stock returns (or equity risk premia) seem to be too high to be explained by investors' consumption optimisation behaviour, implying implausibly high levels of their risk aversion. Shiller (1981) and others have noted that stock prices exhibit excessive volatility, as compared to changes in fundamentals. There are also some indications that markets may be more predictable than the EMH hypothesis suggests (Campbell and Shiller, 1988; Lo and MacKinlay, 1988). Empirical facts, such as large trading volumes, fat tails of returns distribution or persistent stock price volatility, are not well understood either (LeBaron, 2006). And, of course, booms, busts and financial crises – which clearly are salient features of today's economic reality and should be placed really high on economists' research agenda – are in discord with the standard rational representative agent paradigm and the EMH hypothesis.

1.2. The new paradigm – markets as complex agent-based systems

Making strong assumptions in standard financial models may have been about the only way to make theoretical generalisations about the market behaviour. But now that the growing computing power and advancing computational methods have enabled researchers to relax some of those assumptions, economics and finance are witnessing an important paradigm shift towards a behavioural, agent-based approach. Pursuant to this approach, markets are seen as complex dynamical systems consisting of heterogeneous learning, boundedly rational heterogeneous agents (Hommes, 2006; LeBaron, 2006).

Computational study of these dynamical systems of interacting agents is what agent-based computational finance is all about. Let us very briefly discuss the principal attributes of the research object of agent-based financial models. Naturally, at the centre-stage are agents. Agents, in this context, are given quite a broad meaning. According to Tesfatsion (2006), they comprise bundled data and behavioural methods representing an entity in a computationally constructed environment. They can range from active, learning and data-gathering decision-makers (e. g. investors, consumers, workers), their social groupings (e. g. firms, banks, families) and institutions (e. g. markets, regulatory systems) to passive world features such as infrastructure. From the operational point of view, they are similar to objects and object groups in the object-oriented programming,

whereas agent-based models technically are collections of algorithms embodied in those entities termed “agents”. The possibility to develop composite and hierarchical structures of computational agents implies that they can become arbitrarily complex and may greatly surpass their analytical counterparts of standard models in respect of reflecting salient features of the real world entities.

The interdisciplinary nature of the notion of an agent also leads one to the realm of the computer science. Here, an autonomous agent is understood as a system situated in, and part of, an environment, which senses that environment, and acts on it, over time, in pursuit of its own agenda (Franklin and Graesser, 1996). If agents are capable of learning to achieve their goals more efficiently or their population as a whole continuously adapts to be better suited to survive, the artificial intelligence theory comes into play. Learning and adaptation are crucially important in agent-based modelling since the ultimate goal of any economic analysis is to model the actual human intelligent behaviour and its consequences at the individual or aggregate level.

Agents form complex adaptive systems. A system is said to be complex if it is constituted of interacting elements (agents) and exhibits emergent properties, i. e. properties inherent to the system but not necessarily to individual agents. Depending on the complexity of studied phenomena, complex adaptive systems may include reactive agents (capable of reacting in a systematic way to changing environmental conditions), goal-directed agents (reactive and capable of directing some of their actions to achieving their goals) and planning agents (goal-directed and capable of exerting some control over environment). It is important that these systems are self-sufficient or dynamically complete, i. e. they may evolve – without interventions from the modeller – in reaction to exogenous environmental changes or even as a result of merely endogenous interaction of agents[3].

Once agents are put together in a complex system, systemic patterns resulting from the agents’ interaction may be observed and the system’s reaction to exogenous shocks can be analysed. Hence, agent-based financial models basically constitute a simulation tool. This determines the delicate position of agent-based financial modelling among standard scientific inference methods: deduction-based theoretical models (i. e. theoretical generalisation from certain assumptions) and induction-based empirical models (recognition of systematic patterns in the empirical data). Simulation, and agent-based modelling in particular, does not allow one to prove theoretical propositions, nor does it directly measure real world phenomena, so there is always a risk of analysing an artificial world too remote from reality. On the other hand, simulation, just like

[3] For a more thorough discussion of the basic elements and principles of agent-based computational economics and finance see Tesfatsion (2006).

deductive analysis, is based on explicit assumptions, which in many cases are much more realistic than in analytical models. If those assumptions and parameters of exogenous processes are calibrated to match empirical data, then simulation analysis does lend itself to drawing valuable inductive inferences about the real world behaviour. Generally, as Axelrod and Tesfatsion (2006) observe, simulation permits increased understanding of systems through controlled computational experiments. Epstein (2006) also notes the importance of agent-based modelling as a tool for generative explanation. While most economic and financial theory deals with analysis of equilibria, he argues that it is not enough to claim that a system – be it an economy, financial market or other social grouping consisting of rational agents – if put in the Nash equilibrium, stays there. For a fuller understanding of a system's behaviour, for generativists it is important to understand how the local autonomous interactions of atomistic, heterogeneous and boundedly rational agents generate the observed macro-level regularities and how the system reaches, if reaches at all, the equilibrium. Moreover, plausibility of any equilibrium patterns at the macro-level is required by generativists to be confirmed by generating it from suitable microspecifications. As Epstein (1999) puts it, "If you didn't grow it, you didn't explain it". In general, agent-based economic and financial modelling has several primary objectives – the abovementioned empirical understanding of macro-level regularities, normative understanding of potential institutional and policy improvements, and qualitative insight and theory generation through examination of simulated behaviour[4].

Key features of agent-based financial models are well summarised by Epstein (2006). The most important feature and actually the primary reason for departing from standard analytical settings is heterogeneity of agents. Agents may differ in their preferences, skills, decision rules, information sets, levels of wealth, etc., and their characteristics may change over time independently of others. Agent behaviour is generally characterised by bounded rationality, which arises both from limited information and limited computational capacities of agents. Agent interactions are autonomous, i. e. there is no central planner, Walrasian auctioneer or other central controllers, though interaction rules, behavioural norms and institutional settings may be arbitrarily sophisticated. Agent-based models also require an explicit agent interaction network, which may be centralised or decentralised (in which case agents interact locally), and their financial decisions may be influenced by information flows through social networks. Finally, analysis of non-equilibrium dynamics of analysed systems in agent-based modelling is of no lesser importance than studying equilibrium properties.

[4] For extended discussion see Tesfatsion (2006).

Agent-based modelling clearly gives researchers a large degree of much desired flexibility necessary to understand the real world financial market phenomena. Unfortunately, this poses problems too. Much room for manoeuvre implies that agent-based models vary to such an extent that they lack some unifying fundament, which could help to develop this interesting area of research into a solid theory with an established methodology and conventional wisdom about basic building blocks. There are also serious difficulties related to modelling micro-level behaviour. Having made the pretty obvious proposition that the real world investors are less than fully rational, researchers face difficult conceptual issues related to deciding what then governs agents' behaviour and how to model it. Should modelled micro-behaviour match that of human subjects in laboratory experiments? Should modellers deliberately include in their agent-based models behavioural biases and heuristic rules confirmed by experimental data? Should artificial agents apply decision rules that have some substantiation in the theoretical representative-agent models? How learning and expectation formation processes should be modelled? Answers to these questions, of course, depend on the problem at hand but, as stressed by Duffy (2006), in principle they are not yet systemically addressed by researchers of this field. Dealing with these issues, Duffy suggests that external validation of simulation results should not be limited to comparison of aggregate outcomes of simulated and real world phenomena. He advocates careful selection of model parameters based on experiments with human subjects and suggests seeking stronger external validation by comparing simulation results to those obtained in the experiments with humans, both at the micro- and macro-level.

1.3. Concluding summary of Chapter 1

Global economic and financial problems caused by the disorderly unwinding of imbalances that had been accumulating over decades cannot be easily reconciled with the efficient market paradigm based on rational expectations and perfectly rational representative agent assumptions. Some even see the collapse of global consumption as a result of the disregard of intertemporal budget constraints, which is of course at odds with basic principles of consumer (investor) optimising behaviour. In any case, the mainstream financial and economic theory cannot explain current economic developments, give adequate forecasts or policy prescriptions. In this situation there is actually a growing need to look for a replacement for the work-horse financial and macroeconomic models based on heroic assumptions. This chapter advocates the view that agent-based financial modelling, which takes financial markets as complex dynamical systems consisting of interacting heterogeneous agents, can potentially become a

viable alternative. However, significant further interdisciplinary progress is needed before agent-based models can take the centre-stage of financial market modelling.

In this chapter we discussed, at the conceptual level, the need for the proper account of market complexity, agent heterogeneity, bounded rationality and adaptive (though not simplistic) expectations. These features are usually absent in standard financial models but are at the heart of agent-based modelling. Agent-based models use the procedural (computer code) rather than mathematical model description form, which technically enables researchers to easily inquire into many interesting features of systemic behaviour. However, theoretical or empirical foundation of individual agents' behaviour and external validation of model results are two main problem areas, which have not been systemically addressed to date and constitute a serious obstacle to the further progress of agent-based financial modelling.

2

Artificial stock market literature review and discussion of model design issues

Artificial stock market[5] (ASM) modelling is one specific area of the agent-based modelling. This chapter is aimed at providing a brief and selective introduction to the emerging body of literature on the agent-based financial modelling and, more specifically, the artificial stock market (ASM) modelling. To our knowledge, such artificial stock market modelling has not been conducted in the Lithuanian economic literature but in this chapter we briefly touch upon the research conducted by Lithuanian researchers in related fields. In this section we also discuss important model design issues arising when developing ASMs.

2.1. ASM building principles and problems

ASM models are usually developed with the aim of examining emergent properties of financial markets. Some of the most important questions addressed

[5] ASMs are also referred to in the literature as simulated stock markets and agent-based models of stock markets.

by these models relate to the self-organisation of financial markets through actions of boundedly rational agents, explanation of empirical patterns of market dynamics (e. g. price, volatility dynamics) and emergence of effective trading strategies. A stylised process flow chart in a typical ASM model is presented in Figure 2.1. Information flows from the external environment are filtered, as agents form individual perceptions of important aspects of the environment. Agents' properties govern their behaviour, which results in the transformation of individual perceptions into actions. Agent interaction takes place in the market institution, where the market price is set as a result of agent interactions. In the case of models supporting adaptation and learning, agent interaction and the resultant market price dynamics may create environment feedback signals needed for systemic adaptation or individual learning. Through these processes agent properties are altered to maintain an appropriate adaptation to the changing environment.

ASM modellers have to deal with challenging design issues and face modelling tradeoffs. These include (but are not limited to) the choice of agents' preferences and objectives, properties of securities, mechanisms of price determination, expectations formation, evolution and learning algorithms, timing issues and benchmarks.

Since stock markets usually constitute an extremely complex environment, one immediate problem is that any degree of realism imposes huge computational costs and results in the loss of analytical tractability. The problem is most severe in modelling agents' intelligent behaviour. It is further aggravated by the fact that the decision-making processes are unobservable and there is too little theoretical guidance on how to model these processes realistically. For these reasons ASM models usually are highly stylised, and in many cases model settings are kept close to certain benchmarks – standard theoretical rational expectations models or tractable and well understood special cases of agent-based models. This generally strengthens the credibility of ASM models as tools of generative explanation of systemic equilibria derived under strong assumptions, and eases interpretation of simulation results.

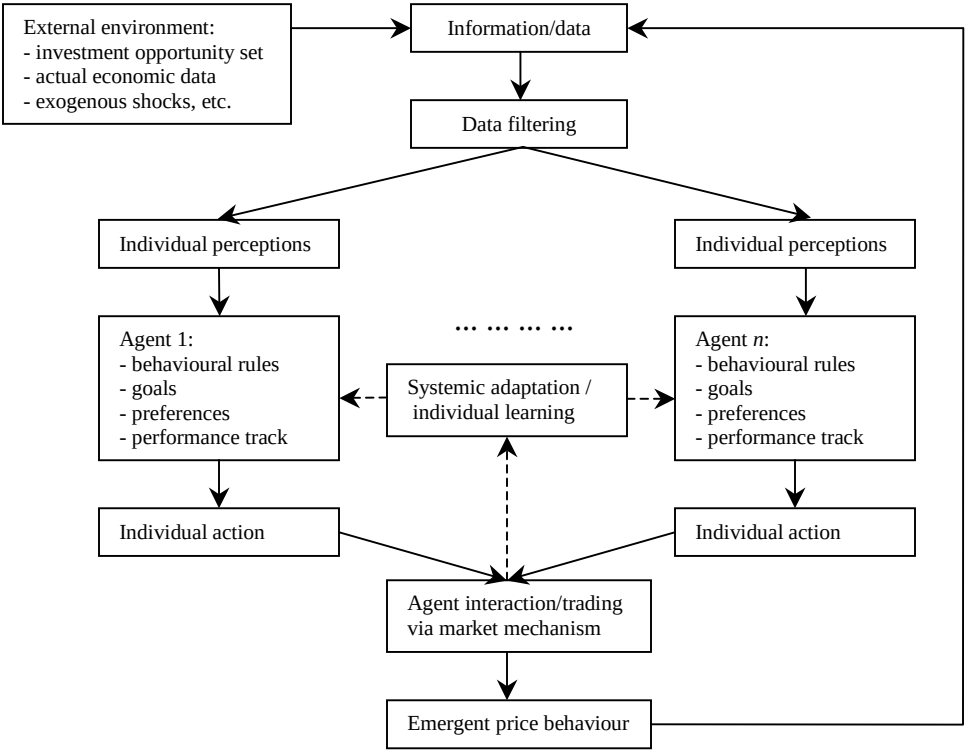


Fig. 2.1. Basic process flow in ASM models

Source: Adapted from van den Bergh, *et al.* (2002).

Agents and their decision-making processes occupy the centre-stage in agent-based models of stock markets and are the main source of diversity of ASMs. The agent design might vary from budget-constrained zero-intelligence agents to sophisticated artificially intelligent decision-making entities. It is worth noting in passing that some agent designs lack dynamic integrity and lasting identity inherent to human subjects, which brings interpretation of these artificial agents closer to competing bundles of strategies rather than to actual investors. Agents usually are given utility functions, and utility levels associated with different strategies are important in driving agents' behaviour or determining their "fitness" in the evolutionary selection process. Agents may derive utility from different sources, e. g., consumption, wealth or returns. A serious limitation but a very natural one, given the complexity of the model environment, is that in most cases agents are myopic in that they care only about one-period utility and do not attempt to carry out dynamic optimisation.

In simplest settings agents' behaviour may be completely random (constrained only by budget constraints to make it economically interesting, as in Gode and Sunder, 1993). Alternatively, they may follow strict decision rules or choose conditional strategies from a (dynamically evolving) bundle of strategies. These rules may be suggested by standard theories or may mimic popular actual investment strategies. The central design question in the ASM models based on artificial intelligence is how agents choose investment strategies and how the pool of available strategies evolves. It should be noted that in almost all models, agents – taken individually – have very limited intelligence, whereas systemic adaptation and strategy improvement mostly take place at the population level.

Many ASM models employ Holland's (1975) genetic algorithm technique to drive the evolution of strategies. In such algorithms, inspired by the theory of biological evolution, strategy pools evolve as a result of the rule crossover, mutation and evolutionary survival of the fittest rules. Alternatively, agents may choose their investment strategies or form their forecasts based on neural network or simple econometric forecasting techniques. Another learning possibility is the Roth and Erev (1995) type stimulus-response learning[6] (discussed, e. g., in Brenner, 2006; Duffy, 2006), which leans on the simple idea that actions yielding larger payoffs tend to be repeated more frequently. More technically rigorous and economically appealing is the reinforcement learning mechanism established in the artificial intelligence literature (see, e. g., Sutton and Barto, 1998). This approach has been hardly ever used in the context of ASM modelling, partly due to some known problems of the application of such algorithms in multi-agent settings. But, in our view, the inability to ensure that the learned strategies are asymptotically optimal should not preclude modellers from taking advantage of these economically intuitive learning algorithms.

Specification of the market setting is another very important ASM design issue. ASM modellers usually simplify the portfolio allocation task, and in most models there are only two types of securities traded – a risky dividend-paying stock and a riskless bond. Moreover, pricing in the bond market is typically shut down by assuming constant interest rates. Pricing of the stock is hence determined by both fundamental factors and the interaction of heterogeneous agents (and dynamics of their expectations), though some features of these determinants may be greatly simplified for analytical or computational purposes. For instance, the dividend process may not be modelled specifically, or dividend can be assumed to be paid out every trading period, which is a highly unrealistic

[6] It is also known in the agent-based modelling literature as reinforcement learning but it should not be confused with the formal reinforcement learning algorithm developed in the artificial intelligence literature.

but quite necessary assumption in the myopic agent environment. Next critical issue in specifying the market setting is the choice of the price determination mechanism. According to LeBaron (2001b, 2006), there are four major classes of price determination mechanisms:

- Gradual price adjustment, in which case individual sell and buy orders (for a given price) are aggregated and in the next trading period the price is gradually shifted by the market-maker in response to excessive supply or demand (the market is almost never in equilibrium);
- Immediate market clearing, whereby the market clearing price is computed from agents' supply and demand functions (the market is always in temporary equilibrium);
- Randomly matched trading, whereby trade takes place between randomly matched agent pairs;
- An order book, which most closely models the actual trading process on order-driven automated stock exchanges.

In this context it is useful to note the problem of trade synchronicity, which is not an issue in standard analytical representative agent models, where there is simply no trade. Actually, the real world traders arrive in the market and make their orders asynchronously, which may lead to strategic intra-period interaction. There are some attempts to build event-driven ASMs instead of ASMs evolving in equal time increments. However, owing to technical and conceptual difficulties, most ASM models assume that trading decisions are taken by all agents simultaneously without having any strategic interaction of this type.

2.2. Brief ASM literature review

Though early attempts to build agent-based financial models were dated a few decades ago, this line of research is arguably still at its early stages of development and is only starting to shape a solid and unified field of the financial theory. Most noteworthy achievements of these models are their abilities to generate the seemingly non-equilibrium properties of financial market dynamics, including clustered volatility of stock returns, heavy tails of returns distribution, bubbles and crashes, and large volumes of trading (Farmer and Geanakoplos, 2008). On the other hand, a down-to-earth explanation of particular episodes of empirical financial developments and prediction of market dynamics (e. g. bubbles and crashes) by using agent-based financial models still seems as a distant future goal. Other weaknesses of these models include vague and sometimes simplistic representation of individual investor behaviour, opaque learning processes, as well as severe restrictions on the number of assets, their payout structures and sets of possible investment strategies.

The need for this alternative modelling of stock market behaviour arose from dissatisfaction with the abovementioned strong assumptions of the standard financial theory, its neglect for simple investment behaviour rules that are often used by finance practitioners and inability of standard models to explain satisfactorily the real world stock price dynamics (e. g. the US stock market crash on 17th October 1987 and, of course, the recent global financial meltdown).

Now let us turn specifically to some well-known models. We divide (somewhat arbitrarily) the models into two broad categories:

- Models based on stochastic, heuristic and standard theory-implied behavioural rules;
- Models with learning agents or evolutionary systemic adaptation.

The latter category is arguably more promising and interesting.

2.2.1. ASM models based on random, heuristic and hard-wired behavioural rules

The first group of ASM models generally investigate whether interaction of heterogeneous agents, who base their decisions on simple deterministic rules or even act in a random manner, might induce stock price movements qualitatively similar to those observed in real stock markets. Agents in these models usually follow simple, "hard-wired" investment rules. In order to generate interesting market dynamics without sacrificing model parsimony and tractability, it is very common to allow just a few investment strategies (these are the so-called few-type models; see LeBaron, 2006). For instance, market participants may be broadly categorised as "fundamentalists", "chartists" and "noise traders". Fundamentalists base their investment decisions on fundamental information about stock dividend potential, chartists rely on technical analysis of time series of stock prices, whereas noise traders may base their investment decisions on erroneous signals about fundamentals, follow aggregate market behaviour or, say, simply behave in a random manner. Though such strategies bear some resemblance to the real world investment behaviour, the problem lies in determining the distribution of different investor types in model population, as this distribution may play a key role in shaping the aggregate market behaviour. Clearly, market developments are influenced by relative popularity of different strategies. Under different circumstances some strategies may become dominant and optimal to follow, hence in some models agents are allowed to switch to different strategies. They can switch to alternative strategies depending on their performance, or underperforming investors may simply exert smaller influence on market developments due to their smaller financial wealth.

The origins of this strand of literature are linked to the few-type models of foreign exchange markets proposed by Frankel and Froot (1988), Kirman (1991), De Grauwe, *et al.* (1993) and others. A prominent early example of the few-type model of a stock market is Kim and Markowitz (1989). In their stylised model there are two types of agents that pursue either the portfolio rebalancing strategy or the portfolio insurance strategy. Rebalancers aim at keeping a constant fraction of their assets in a risky stock, while portfolio insurers try to ensure the minimum level of wealth by defensively selling some of the stock holdings if the minimum threshold approaches. The rebalancing strategy works as a market stabilising force (a stock price decline spurs stock buying), whereas the portfolio insurance strategy amplifies market fluctuations (a stock price decline prompts stock selling). With simple expectation formation and decision rules and with the uncertainty induced by monetary shocks, the model shows that some investment strategies, namely the abovementioned portfolio insurance strategy, may have a sizeable destabilising effect on the market and can be partly responsible for market crashes. Later models developed along this line of research are much more detail-rich but they are still aimed at giving a plausible explanation of complicated empirical market dynamics, which is not quite consistent with standard financial models. In detailed models of Lux (1995), Lux and Marchesi (1999, 2000) much of market dynamics is attributed to agents' endogenous switching to different trading strategies depending on the prevailing majority opinion. Another popular idea in the ASMs is that some "smart money" traders possess superior information and are able to form (boundedly) rational expectations, while others are noise traders (Shiller, 1984; DeLong, *et al.* 1990). Some models consider the choice between costly optimisation and cheap imitation strategies (see e. g. Sethi and Franke, 1995). In several other models, artificial agents follow investment strategies based on standard theoretical principles, such as standard mean-variance optimisation (Jacobs, *et al.*, 2004; Sharpe, 2007). Finally, it should be noted that contributions from econophysicists in this area of research are very significant – see Samanidou, *et al.* (2007) for a review. A lot of agent-based financial models actually are a product of economists' close collaboration with physicists, drawing on the experience of the latter in studying systemic behaviour resulting from complex interaction of atomistic particles.

2.3.1. ASM models based on intelligent adaptation

Rational-expectations representative-agent models examine optimal investment strategies and asset pricing in the equilibrium. ASM models based on hard-wired investment strategies mostly deal with emergent properties of stock markets. All of this is important in ASM models based on intelligent adaptation. In addition

to this, these ASM models also help to explain how investors may come up with good strategies, examine their stability, and whether such artificial markets can generate equilibria derived from standard models with restrictive assumptions. Basic features of the two broad categories of the ASM models are compared in Table 2.1. It should be noted, however, that the dividing line between these model categories is blurred and some models may share some elements with the other group.

Table 2.1. Comparison of the two broad categories of ASM models

	Models with hard-wired behavioural rules	Models supporting systemic adaptation
Possible model purposes	Provide possible explanation for anomalous market movements, produce outcomes similar to empirical patterns, analyse how and if market reaches equilibrium, compare to benchmark theoretical models	The same and examine emergence of new investment strategies, analyse investment strategy evolution and stability
Agent properties	Mostly few-type, hard-wired rules	Often many-type, dynamically changing rules
Systemic adaptation	No or little	Yes, through evolutionary selection or individual learning
Individual learning	No	Possible, based on neural nets, econometric forecasting, stimulus-response, reinforcement learning, etc.
Agent type distribution	Mostly modeller-determined	Initially modeller-determined but changes as a result of systemic adaptation
Model outcomes	Emergent systemic properties, market price dynamics	Emergent systemic properties, market price dynamics, agent type distribution dynamics, evolution of strategies, possibly new strategies

Source: formed by the author.

Dynamically evolving and improving strategies is a remarkable feature of these agent-based financial market models. It greatly reduces reliance of modelled market behaviour on arbitrarily chosen investment strategies and increases model realism. Participants of real world financial markets act in a highly uncertain environment and most of them try to adapt to changing environmental conditions and learn to improve their strategies. Learning and adaptation mechanisms in ASM models may still be very far from anything that realistically describes genuine human learning but, in any case, attempts to model this crucial feature of investment behaviour constitute a qualitatively different approach to financial modelling. In most ASMs, artificial intelligence techniques, most notably evolutionary algorithms and neural networks, are favoured over simplistic adaptive rules. Since intelligent adaptation implies choosing among many different strategies, as well as creating new ones, there is usually a large and evolving ecology of investment strategies in these models, and they are therefore sometimes referred to as the "many-type" models.

A detailed review of the related ASM literature is provided by LeBaron (2006). Here we only briefly mention some of the more popular models. A model proposed by Lettau (1997) is one of the early attempts to examine whether a population of heterogeneous agents are able to learn optimal investment strategies in a very simple market setting, for which the analytical solution is known. In his model, agents are endowed with myopic preferences and have to decide what fraction of their wealth to invest in the risky asset with the exogenously determined random return. As is very common in ASM models, evolutionary systemic adaptation is ensured by application of the genetic algorithm. Even though individual agents actually have no intelligence, fitter individuals (i. e. those whose strategies give higher levels of utility) have better chances of survival, and this evolutionary selection leads to near-optimal strategies over generations in this simple model. Routledge (1999, 2001) examines adaptive learning in financial markets in a more complex setting, namely, a version of Grossman and Stiglitz's (1980) model of heterogeneous information about future dividends and signal extraction. He presents an analytic framework for adaptive learning via imitation of better-informed agents and shows that the rational expectations equilibrium is broadly supported by adaptation modelled with genetic algorithm.

Another interesting market setup is proposed by Beltratti and Margarita (1993). In this model, agents apply artificial neural network techniques to forecast stock prices from historical data, and trading may take place between randomly matched individuals with differing expectations. A noteworthy feature is that agents may choose to apply either a sophisticated neural network (with more hidden nodes, or explanatory variables) at a higher cost or a cheaper naïve network, and this corresponds to the real world fact that sophisticated investment

is a costly endeavour. Interestingly, in some settings the naïve investors gain ground, once markets settle down and additional benefits from sophisticated forecasting do not cover its cost.

One of the most famous ASMs is the Santa Fe artificial stock market model developed by Arthur, *et al.* (1997), also described in LeBaron, *et al.* (1999) and LeBaron (2006). This model is aimed at exploring the evolution and coexistence of a pool of strategies that compete with each other in the genetic algorithm environment and drive the market toward some informational efficiency. In this model there are two securities: a risky dividend-paying stock and a riskless bond that offers the constant interest rate. Heterogeneous agents are myopic[7] and have constant absolute risk aversion preferences. They try to forecast a next period's stock price by applying simple "condition-forecast" rules and plug the forecast in their (induced) asset demand functions. The equilibrium price is determined by the auctioneer by balancing aggregate demand for shares with fixed supply. Forecast adaptation in this model is based on modified Holland's (1975) "condition-action" genetic classifier system. Each agent is given a set of rules mapping states of the world (such as the relative size of price/dividend ratio or a stock price relative to its moving average) to forecasts (which are a linear combination of the current stock price and dividend). The rules endogenously evolve as a result of cross-over, mutation and selection. The authors of the Santa Fe ASM examine convergence of the market price to the homogenous rational expectations equilibrium, and show that this is the case for certain parameter settings corresponding to the "slow learning" situation. The model is able to generate some statistical features of price dynamics qualitatively similar to stylized facts about the real world financial markets, though no attempt is made to quantitatively line them up with actual financial data or examine the realism of assumptions about dividend processes (LeBaron, 2006). The model served as a platform for a number of further extensions (Joshi, et al., 2000; Tay and Linn, 2001; Chen and Yeh, 2002a).

In an interesting model, LeBaron (2001a) examines how investors' heterogeneous time horizons (i. e. different information sets, upon which agents choose decision rules) affect evolution of the market, its convergence to the known homogenous rational expectations equilibrium and domination of different types of investors. The model is in some respects similar to the Santa Fe model but has some notable original features. The learning mechanism in this model is an interesting combination of the neural network technique and the evolutionary search mechanism. In contrast to most other models, agents learn portfolio allocation decisions rather than explicitly form price expectations. Also, agents have access to a public – rather than private – pool of investment

[7] Agents care just about one period and do not dynamically optimise.

strategies, which are based on simple feedforward neural networks. Agents evaluate these strategies by feeding data series of heterogeneous length into these networks, and this forms the basis of their heterogeneity. Furthermore, the neural networks are evolved by applying mutation, crossover, weight reassignment and rule removal operations, and some of the worst-performing individuals are also replaced by new agents. One of model's main findings is that short-memory (and short investment horizon) investors are not driven out of the market, and acting as some market volatility generators, they hinder attaining the rational expectations equilibrium.

2.4. Related research in Lithuanian scientific literature

Lithuanian researchers have been interested in stock market analysis and modelling of investment strategies for more than a decade but the specific area of artificial stock market modelling has not been systematically researched and, to our knowledge, no full-fledged artificial stock market models have been developed by Lithuanian researchers. More specifically, most studies conducted by Lithuanian scholars focus on one of these areas: research of investment decision making, analysis of emergent properties of stock markets, and analysis of empirical relationship between market dynamics and macroeconomic variables.

Studies of investment strategies are conceptually most closely related to our research. Similarities mostly lie in the recognition of need to step aside from the neoclassical financial theories and possibly employ artificial intelligence techniques to govern investment behaviour. Most notable differences are that in the majority of such models developed by Lithuanian researchers investment strategies are based on population-level adaptation and are applied against real financial data, whereas we emphasise individual adaptation, economically-based individual behaviour and endogenous interaction of artificial agents[8].

We briefly mention some of the more important studies of investment strategies. 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. These same authors

[8] This considerably complicates modelling because it necessitates introducing agent interaction framework, more complicated individual behavioural principles, etc.

attempt at incorporating both fundamental and technical analysis in strategy building (Simutis and Masteika, 2006). Rutkauskas (2008) implements 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.

Turning to the analysis of emergent stock market properties, some authors analyse the efficiency of the Lithuanian stock market. For instance, Kvedaras and Basdevant (2004) analyse time-varying autocorrelation of stock returns and find evidence for a weak form of informational market efficiency. On the other hand, Dubauskas and Teresienė (2005) find that stock returns possess properties of kurtosis, skewness and nonnormality, and claim that this helps to forecast tendencies of stock market dynamics. Girdzijauskas and Štreimikienė (2008), Girdzijauskas, *et al.* (2009) aim at explaining formation of (very long-term) economic and stock market bubbles with their logistic theory of capital, which puts a strong emphasis on physical capacity limitations to existing capital.

Several studies of Lithuanian researchers aim at examining empirical relationship between stock market performance and macroeconomic indicators. Pilinkus (2009) performs Granger causality tests to identify leading, lagging and Granger-causing macroeconomic indicators in relation to stock returns. Teresienė, *et al.* (2008) focus in particular on the impact of inflation and interest rates on stock returns. The author of this dissertation also proposed several econometric models to model macroeconomic implications of the financial bubble (Ramanauskas, 2007, Ramanauskas and Jakaitienė, 2007).

2.5. Concluding summary of Chapter 2

This chapter focuses on the artificial stock market modelling. In the brief literature review two broad categories of models are presented: (*i*) models based on stochastic, heuristic and standard theory-implied behavioural rules and (*ii*) models with learning agents or evolutionary systemic adaptation. The first group of models generally emphasise the role of agent heterogeneity in determining complex market behaviour and emergent systemic properties, whereas the second group of models concentrate on the generation of good investment strategies and the macro-level implications of intelligently adapted investment strategies.

To our knowledge, no full-fledged artificial stock market models have been developed by Lithuanian researchers. However, in this chapter we give some references to somewhat related studies conducted by Lithuanian scholars. They

mostly fall into one of these areas: research of investment decision making, analysis of emergent properties of stock markets, and analysis of empirical relationship between market dynamics and macroeconomic variables.

3

Reinforcement learning in the context of artificial stock market modelling

Standard financial and economic theories assume that economic decision makers know the underlying model of the world, in which they act. This allows them to deductively know the consequences of their actions (or the probability distribution of possible outcomes). In contrast, if the model of the world is not known, agents either have to act according to simple adaptive rules or to attempt to know the underlying model through interaction with the environment.

In the models that we develop agents are realistically assumed to be unaware of the underlying model of the economy but instead possess cognitive abilities and try to develop coherent behavioural rules. Their learning is formalised by the machine learning methodology. More specifically, they resort to one particular reinforcement learning algorithm, namely, the gradient descent Q-learning algorithm. In this chapter we provide some economic rationale for choosing this particular learning method, present some basic principles of reinforcement learning and briefly discuss application of these methods for economic problems.

3.1. Economic rationale for choosing reinforcement learning method

Expectations regarding prospects of a particular stock or a stock index play a crucial role in financial modelling because investment decision making relies – explicitly or implicitly – on agents' ability to make a priori assessments of possible consequences of their actions. If one is willing to accept the obvious empirical fact of a large variety of market expectations, or if one is aimed at explaining how this diversity comes about, then it is necessary to abandon the rational expectations assumption, which allowed to them circumvent these issues in standard models. According to Arthur (1995), rational expectations equilibrium is a special and in many cases not robust state of reality, whereby individual expectations induce actions that aggregatively create a world that validates them as predictions.

Abolition of the rational expectations assumption immediately implies that agents no longer know the model of environment. It therefore becomes necessary to assume that either agents' behaviour is governed by simple adaptive rules or they can change and calibrate these rules during the process of learning. In this study we are, of course, more interested in the latter alternative.

How do investors behave in realistic, out-of-equilibrium situations? Needless to say, realistic modelling of expectations formation poses a great challenge and it is essentially a grey area of the financial theory but a few observations about investor behaviour in the highly uncertain environment can be made. In such an environment it seems perfectly sensible for investors to act adaptively and follow inductive reasoning, or in other words, form simple forecasting models, test them and update them depending on their performance. Hence, investors should constantly learn from interaction with environment.

Their learning is not a supervised process because due to model uncertainty generally there is no way of knowing the true intrinsic value of a stock even in retrospect. For instance, if an investor observes a stock price realisation, which is different from what he had expected, he generally cannot know whether this deviation is attributable to his misperception of fundamentals, unpredictable shocks to fundamentals, complex interaction of market participants or other factors. However, even without knowing retrospectively what the "correct" expectations and actions should have been, investors can judge about adequacy of their beliefs and actions by reinforcement signals that they receive from interaction with the environment. Possible reinforcement signals include their performance relative to the market, long-term portfolio returns, utility from consumed earnings, etc.

Another important aspect is that in order to better adapt in the uncertain environment investors have to both exploit their accumulated experience and explore seemingly suboptimal actions.

If the above-described investment behaviour is deemed an adequate description of how boundedly rational investors actually behave, reinforcement learning methods developed in the artificial intelligence literature seem to be conceptually suitable for modelling investor behaviour (though there are some problems with technical implementation).

Inspired by the psychology literature, reinforcement learning is the sub-area of the machine learning, and in its core lies agents' interaction with environment in pursuit of highest long-term rewards. It can potentially be of a particular interest to economists because standard theoretical economic agents' behaviour is often guided by very similar principles. In standard economics and finance, agents choose action plans that ensure maximisation of life-time utility (long-term rewards), and that is exactly what reinforcement learning agents seek – the main difference being that the latter do not know the underlying model of the economy. The well-established link between the basic reinforcement learning algorithm and dynamic programming, as well as proven ability of some reinforcement learning algorithms to achieve (under certain conditions) convergence to optimal policies are especially attractive features of this methodology, from the economists' viewpoint. The reinforcement learning agents are also well-positioned to solve the temporal credit assignment problem, i. e. determine strategic actions that enable them to reach their ultimate goals even though those actions may not be attractive in the short-term. For economists, it is a great advantage over simpler forms of adaptive learning.

3.2. Basic principles of reinforcement learning

Reinforcement learning addresses the question of how an autonomous agent that senses and acts in its environment can learn to choose optimal actions to achieve its goals (Mitchell, 1997, p. 367). More specifically, by taking actions in an environment and obtaining associated rewards, a reinforcement learning agent tries to find optimal policies, which maximise long-term rewards, and the process of improvement of agent policies is the central target for reinforcement learning methods. A good introduction to the reinforcement learning techniques may be found in Sutton and Barto (1998), Bertsekas and Tsitsiklis (1996) and Mitchell's (1997) books, and some broad overview of reinforcement learning models is given in Kaelbling, Littman and Moore (1996) survey. In this subsection we present briefly some basic principles of the reinforcement

learning methodology with a special emphasis on Watkins' Q-learning algorithm, as it forms the basis of agent behaviour in our ASM model.

The iterative sequence of agent's interaction with environment is as follows. At time *t*, the agent observes environment state s_t and acts according to its action policy to produce action a_t. In the next time step it receives numerical reward signal r_{t+1} from the environment and observes new state s_{t+1}. Finally it is ready to update its policies (if necessary) and take new action a_{t+1}. In the reinforcement learning problems it is also assumed that environment possesses the Markov property, i. e. all relevant information about possible future development of environment is encapsulated in the information about the current state and action. More formally,

$$\begin{aligned}\Pr\{s_{t+1} = s', r_{t+1} = r \mid s_t, a_t, r_t, s_{t-1}, a_{t-1}, r_{t-1}, ..., r_1, s_0, a_0\} = \\ = \Pr\{s_{t+1} = s', r_{t+1} = r \mid s_t, a_t\}.\end{aligned} \tag{3.1}$$

If condition (3.1) holds, such reinforcement learning task is called a Markov decision process. To completely specify the environment dynamics for a Markov decision process, it suffices to define state transition probabilities and expected rewards. State transition probabilities constitute a distribution of probabilities of each possible next state s', given any current state *s* and action *a*:

$$P_{ss'}^{a} = \Pr\{s_{t+1} = s' \mid s_t = s, a_t = a\}. \tag{3.2}$$

Notably, in a general case, state transition probabilities are not known to the reinforcement learning agent but can be inferred from interaction with environment. The expected next reward is

$$R_{ss'}^{a} = E\{r_{t+1} \mid s_t = s, a_t = a, s_{t+1} = s'\}. \tag{3.3}$$

As was mentioned above, learning is understood in this context as an attempt to find optimal policies. Here, a policy is defined as a mapping from each state *s* and action *a* to the probability $\pi(s,a)$ of taking action *a* when in state *s* (if a policy is deterministic, then it is simply a set of deterministic rules describing how to behave in each state). For the further elaboration of the reinforcement learning task, the notion of value functions should be introduced. The state-value function for policy π is defined as the expected discounted cumulated reward conditional on state *s* and policy π:

$$V^{\pi}(s) = E_{\pi}\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s\right\}, \tag{3.4}$$

where E_π denotes the expectation given that the agent sticks to its policy π, and γ is a discounting parameter. It proves very useful to define also the value of taking action *a* in state *s* under policy π. The action-value function is given by

$$Q^\pi(s,a) = E_\pi\left\{\sum_{k=0}^{\infty}\gamma^k r_{t+k+1} \mid s_t = s, a_t = a\right\}. \tag{3.5}$$

It is obvious that both value functions possess the Bellman property, i. e. they must be dynamically consistent. For instance, it follows from equation (3.4) that

$$V^\pi(s) = E_\pi\left\{r_{t+1} + \gamma V^\pi(s_{t+1}) \mid s_t = s\right\}. \tag{3.6}$$

Since condition (3.6) holds for all value functions, it also holds for optimal value functions, i. e. those associated with optimal policies[9]. This leads directly to Bellman optimality equations for the state-value function

$$V^*(s) = \max_a E\left\{r_{t+1} + \gamma V(s_{t+1}) \mid s_t = s, a_t = a\right\} \quad \text{for all } s \tag{3.7}$$

and for the action-value function

$$Q^*(s,a) = E\{r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') \mid s_t = s, a_t = a\} \quad \text{for all } s \tag{3.8}$$

The most prominent feature of Bellman optimality equations is that they actually rearrange the multi-period optimisation problem into a problem consisting of a set of difference equations (one for each state). Notably, if value functions are known, it becomes very easy to find optimal policies. Equation (3.7) implies that in any state *s* it suffices to take the greedy action (that is, concerned with only one period ahead) that maximises the expected sum of the immediate reward and the (discounted) next state-value[10]. It is even simpler if the problem is expressed in terms of known action-value functions – from equation (3.8) it follows that action a' taken in state s_{t+1} will be optimal if it maximises the associated expected action-value function. To put differently, it is optimal to take actions that simply maximise each period's Q-function value (such actions are sometimes called Q-greedy actions).

---

[9] Optimal policies are defined as policies that maximise state values V^π in all states.

[10] Notice that expectations are no longer conditioned on specific policies in equations (3.7) and (3.8).

The big question is, of course, how to find optimal value functions. One of the ways to do this is to apply dynamic programming, which also provides the foundation for reinforcement learning methods. The basic idea is to apply some iterative procedure aimed at evaluating current policies and gradually improving them until they converge to optimal policies. More specifically, the so-called generalised policy iteration consists of two interacting processes: (*i*) policy evaluation, which is the process of finding the value function for an arbitrary policy, and (*ii*) policy improvement, whereby policies are improved by making them greedy with respect to the current value function.

The policy evaluation procedure uses Bellman equation (3.6) as an update rule:

$$V_{k+1}(s) = E_\pi \{ r_{t+1} + \gamma V_k(s_{t+1}) \mid s_t = s \}, \tag{3.9}$$

where V_k denotes the *k*-th approximation of the state-value function (V_0 is chosen arbitrarily). It can be shown that estimate V_k converges to true policy V^π as *k* converges to infinity. Each iteration is a sweep through all states – the value of every single state is backed up using equation (3.9).

The policy improvement step is closely linked to Bellman optimality equation (3.7). It can be shown that for every state *s*, the policy can be improved by taking action that maximises the immediate action value or, in other words, looks best in the short term (examining only one period ahead):

$$\pi^* = \arg\max_a E\{ r_{t+1} + \gamma V(s_{t+1}) \mid s_t = s, a_t = a \}. \tag{3.10}$$

The two procedures, given in equations (3.9) and (3.10), are implemented alternately in each iteration, and the iterative process continues until state values and associated policies stabilise, which is when they become optimal. The problem with the dynamic programming is that in order to implement these back-up sweeps, state transition probabilities $P^a_{ss'}$ and expected rewards $R^a_{ss'}$ (see equations (3.2) and (3.3)) must be known, and it is very rarely the case in practice.

A natural way to overcome the problem of incomplete information is to use sample estimates instead of expectations. This is exactly what is done in two broad classes of reinforcement learning, namely, Monte Carlo methods and temporal difference models of learning. In the remainder of this section we present just one specific temporal difference learning method devised by Watkins (1989), also known as the Q-learning. This method's principal back-up rule is closely related to Bellman optimality equation (8) and is of the following form:

$$Q(s_t, a_t) \leftarrow (1-\alpha) \underbrace{\cdot Q(s_t, a_t)}_{\text{Old estimate of } Q(s_t, a_t)} + \alpha \underbrace{(r_{t+1} + \gamma \max_a Q(s_{t+1}, a))}_{\text{New estimate of } Q(s_t, a_t)}. \quad (11)$$

There are two differences from the dynamic programming update rule based on the Bellman optimality condition. First, as was already mentioned, the expectations operator is gone – the actual realised reward and actual action value from the look-up table are used instead of the expected reward and expected Q-value, respectively. Second, the Q-value in the look-up table is not directly replaced with its new estimate but is rather averaged with the previous estimate (which provides needed additional stability for the convergence to the correct Q function). The speed of learning, of course, depends on the learning parameter α – higher values of the learning parameter ensure faster learning. Higher values of α may be useful at the beginning of the learning process as the learning starts from arbitrary policies, or in nonstationary environment where the reinforcement learning agent needs to adapt faster and more flexibly.

It is shown that under quite general conditions the update rule (11) guarantees convergence of the action-value function to the optimal Q-function, provided all state-action pairs are visited infinitely many times. The latter condition is needed to avoid early convergence to suboptimal policies. It requires that the learning agent continues to explore the environment by occasionally taking seemingly suboptimal values so as to ensure that all actions in all states are sufficiently explored. Hence, the Q-learning agent follows the Q-greedy policy most of the time but sometimes (e. g. with prespecified probability ε) takes an exploratory action, which may be completely random or oriented towards more efficient exploration. Such a behavioural policy is usually called ε-greedy.

Having discussed the basic principles of the Q-learning agent's behaviour, now it is possible to describe its behaviour in the procedural form – see the pseudo-code in Figure 3.1. Unfortunately, this simple algorithm can be rarely applied in practice. The reason is that it requires representation of the Q-function as a table with one entry for each state-action pair. This is not possible if the state space is continuous. Even in discrete real-world problems – and especially in the problem of investment behaviour modelling – the size of the Q-table and the computational burden associated with back-up operations are basically unmanageable. This implies that usually it is impossible for the Q-learning agent to fully explore the state space and it is necessary to generalise its prior experience to unfamiliar, but qualitatively similar state-action pairs that are of interest. Such generalisation is also called structural credit assignment – another important feature of the reinforcement learning.

Initialise *Q(s,a)*, *s* arbitrarily
Repeat:
Choose *a* using policy derived from *Q* (e. g. ε -greedy)
Take action *a*, observe *r*, *s'*
$$Q(s,a) \leftarrow (1-\alpha)\cdot Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a')]$$
$$s \leftarrow s'$$
until convergence is achieved or process is terminated

Fig. 3.1. Basic Q-learning algorithm

Source: adapted from Sutton and Barto (1998).

There are a number of readily available methods for experience generalisation. In our model we use the standard linear gradient-descent function approximation for the Q-function, which we now describe briefly.

The idea of the linear approximation procedure is to replace the representation of the Q-function as a look-up table with some linear function and iteratively update its parameters instead of updating Q-values for every single state. Hence, the estimate of the action value function is replaced by the following linear approximation:

$$Q_t(s,a) \approx \sum_{i=1}^{n} \Theta_t(i,a)\vec{\phi}_s(i), \quad \text{for all } a. \tag{3.12}$$

Here $\vec{\phi}_s$ is the $n \times 1$ vector of state features, i. e. arbitrarily chosen variables that reflect the distinctive features of a given state. Matrix Θ_t is the $n \times m$ parameter containing parameters associated with *n* state features for each of *m* possible actions. For more intuitive exposition it is convenient to work with column vectors of this matrix.

The gradient-descent methods seek to gradually adjust the current approximation of the Q-value toward its new estimate and the step size is proportional to the negative gradient of some measure of current deviation (e. g. mean squared error). More specifically, for a given action *a*, the parameter vector $\vec{\theta}_t$ can be updated as follows:

$$\vec{\theta}_{t+1} = \vec{\theta}_t - \frac{1}{2}\alpha \nabla_{\vec{\theta}_t}[v_t - Q_t(s_t,a_t)]^2, \qquad \text{for all } a, \tag{3.13}$$

where v_t is the new approximation of the action-value function and serves as a training example for the parameter update, and $\nabla_{\vec{\theta}_t} f(\vec{\theta}_t)$ is the gradient of this example's squared error, i. e. the column vector of partial derivatives of function *f* with respect to elements of $\vec{\theta}_t$. By taking the derivatives in equation (3.13), one gets

$$\vec{\theta}_{t+1} = \vec{\theta}_t + \alpha[v_t - Q_t(s_t, a_t)]\vec{\phi}_s, \quad \text{for all } a. \tag{3.14}$$

The new sample estimate of the action-value function, v_t, is obtained similarly to the basic Q-learning algorithm (see equations (3.8) and (3.11)). The parameter update equation (3.14) thus becomes

$$\vec{\theta}_{t+1} = \vec{\theta}_t + \alpha[r_{t+1} + \gamma \max_a Q_t(s_{t+1}, a) - Q_t(s_t, a_t)]\vec{\phi}_s, \quad \text{for all } a. \tag{3.15}$$

This equation forms the basis of the Q-learning algorithm, which is applied by artificial agents in our model when forming expectations about the intrinsic stock value. The detailed procedural form of the algorithm is given in Figure 3.2.

Initialise Θ, ϕ_s, s, a arbitrarily
Repeat:
  Take action *a*, observe *r, s'*
  $\delta \leftarrow r - \vec{\theta}_a^T \vec{\phi}_s$
  For all actions *a*:
    $Q(s', a) = \vec{\theta}_a^T \vec{\phi}_{s'}$
  $\delta \leftarrow \delta + \gamma \max_a Q(s', a)$
  $\vec{\theta}_a \leftarrow \vec{\theta}_a + \alpha\delta\vec{\phi}_s$
  With probability $1 - \varepsilon$:
    $a \leftarrow \arg\max_a Q(s', a)$
  else:
    Choose a randomly
  $s \leftarrow s'$
until convergence is achieved or process is terminated

Fig. 3.2. Gradient-descent approximation Q-learning algorithm

Source: Adapted from Sutton and Barto (1998).

The gradient-descent Q-learning is the so-called off-policy control method, as the value function backup procedure uses the highest Q-value of the resultant state, $\max_a Q(s', a)$, rather than the one associated with the current policy, $Q(s', a')$. Unfortunately, convergence to the optimal solution or its vicinity is not guaranteed for the off-policy methods. Nevertheless Sutton and Barto (1998) suggest that it may be possible to guarantee convergence for the Q-learning algorithm when the Q-function estimation policy and the action policy are sufficiently close to each other, which is the case if the ε-greedy policy is followed. There is also evidence that these methods give good practical performance despite the lack of theoretical guarantees of convergence to optimal policies (Tesauro and Kephart, 2002).

3.3. Potential for application of reinforcement learning technique in economics and finance

Another, potentially more severe problem with practical application of standard reinforcement learning methods for interesting economic and financial problems is that convergence requirements include stationary environment, fully observable states and the single-agent setting. In other words, the reinforcement learning agent is capable of learning to effectively adapt in the well-defined stationary environment but, naturally, simple adaptive learning cannot guarantee optimal behaviour once the multi-agent setting brings in strong non-stationarity and strategic interaction among agents.

In modelling the financial, or any other market as a complex adaptive system consisting of a large number of interacting reinforcement learning agents, one must face the issue of whether standard reinforcement learning techniques can adequately govern agents' behaviour. Existing financial research provides little guidance in this respect because, to our knowledge, no well-known multi-agent stock market models based on reinforcement learning have been developed so far.

The reinforcement learning literature, however, provides some evidence that using the single-agent Q-learning algorithm in the multi-agent setting quite often leads to either exactly or approximately optimal policies (Tesauro, 2002). For instance, Tesauro and Kephart (2002) show that in a stylised two-seller market price-setting policies derived by using the standard Q-learning algorithm outperform some fixed and myopic policies and, in some settings, convergence to optimal policies is achieved. It should also be noted that in order to improve performance in multi-agent settings different extensions to the standard reinforcement learning algorithms have been proposed. They are mainly applied

in two-player games, and they take into account the opponent's estimated strategies (e. g. Littman's (1994) Minimax-Q algorithm, Tesauro's (2004) Hyper-Q algorithm or the Nash-Q algorithm developed by Hu and Wellman, 2003). Alternatively, agents may adapt learning rates according to the current performance, as in the WoLF (Win or Learn Fast) algorithm developed by Bowling and Veloso (2001).

It should be noted that reinforcement learning methods have been quite successfully applied in various portfolio management problems. One of the early applications is Neuneier's (1996) model, which employs the Q-learning in combination with the neural network as a value function approximator for optimal currency allocation in a simple two-currency, risk-neutral setting. Moody and Saffell (2001) apply direct reinforcement learning to optimise risk-adjusted investment returns for intra-day currency and stock-index trading. Van Roy (1999) uses the temporal difference learning algorithm for valuing financial options and optimising investment portfolio. Reinforcement learning methods of portfolio management are gradually gaining popularity among practitioners but theoretical literature remains relatively scarce and further research is needed to unleash the potential of this approach.

In our model, strategic interaction among agents is limited as agents interact in a competitive manner via the centralised exchange, where they participate in double auctions and take decisions simultaneously. The problem is arguably alleviated by the fact that the number of agents is quite large because what matters for any specific agent is the relatively stable distribution of all other agents' actions (buy or sell orders) rather than individual actions per se. Hence, the average market price and other trading statistics can be seen as some summarising functions of multi-agent interaction and this is taken into account when individual decisions are made. Additional stability of the learning processes is guaranteed by combining the Q-learning algorithm with some evolutionary adaptation principles that are described in the following chapter.

3.4. Concluding summary of Chapter 3

Intelligent adaptation in the highly uncertain environment can arguably be key to understanding actual financial market behaviour. One of the possibilities is to resort to the artificial intelligence literature for specific algorithms of adaptation and learning. In particular, we find reinforcement learning algorithms economically very appealing, though there are certain problems with their practical implementation. Without knowing the true model of reality, reinforcement learning agents learn from interaction with environment and adjust their strategies so that they attain maximum long-term reinforcement

(utility) from the environment. Similarly, investors seek to reach their long-term objectives in highly uncertain environment. We continue to explore these ideas further by applying them in our artificial stock market models.

4

Building artificial stock market populated by reinforcement-learning agents

In this chapter 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. The main aim of the proposed model is to offer, implement and test some new ideas that could lay ground for a robust framework for 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 existing agent-based financial models. However, the lack of the widely accepted fundament in this area of modelling necessitates the individual and largely independent approach.

By conducting simulation experiments in this model, we are primarily interested in examining market self-regulation abilities and the congruence

between the market price of the stock and its fundamentals (the market efficiency issue). We also discuss the importance of intelligent individual behaviour and interaction at the population level for market efficiency and functioning and explore the relationship between stock prices and market liquidity.

4.1. Description of the ASM 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.

One of the more interesting features of the present model is a relatively detailed modelling of decision processes. Importantly, agents in this model exhibit economically appealing and forward-looking behaviour, which is based on reinforcement learning and evolutionary change. More specifically, we follow the view, advocated by Arthur (1995) and other proponents of agent-based modelling, that investors exhibit inductive behaviour. Agents do not know the model of the world and cannot deduce optimal actions from optimising conditions but rather explore the environment and try to adapt to it. We resort to the machine learning literature and apply one of the reinforcement learning algorithms, namely the Q-learning algorithm devised by Watkins (1989). It has some highly attractive features from an economist's viewpoint but there are known problems with its practical implementation. To our knowledge, this is

one of the first attempts to incorporate Q-learning algorithm into the ASM models.

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, agents in this ASM realistically 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, skill level and endowment. 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 interesting generative explanation of prolonged periods of over- and under-valuation. In this section we present the architecture of the artificial stock market in detail.

4.1.1. 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 population. We can summarise agents’ basic behavioural principles 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 against the backdrop 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.

Table 4.1. Main building blocks of the ASM model

Forming individual forecasts of exogenously generated dividends Based on: • Exponential moving average • Adjustment as a result of reinforcement learning (agents seek to minimise forecast errors)
Making individual estimates of fundamental stock value and its reservation price Based on: • Discounted expected dividend flows • Adjustment as a result of reinforcement learning (agents seek to maximise portfolio returns)
Making individual trading decisions Based on: • Private estimates of fundamentals, • Maximisation of expected individual wealth at the end of a trading period • Publicly announced estimated probabilities of successful trades for given prices
Carrying out trades via the centralised exchange and collecting trading statistics Based on: • Double auction system • Simultaneous submission of trade orders and random queuing of individual orders
Learning to forecast dividends and learning about fundamental stock value Based on: • Standard Q-learning with linear gradient-descent approximation
Augmenting learning processes by specific interaction among agents (optional) Based on: • Successful strategy imitation • Evolutionary selection and resultant prevalence of successful investment strategies • Noise trading behaviour

For the ease of exposition, it is useful to break the model into a set of economically meaningful processes, though some of them are inter-related in

more complex ways. The general structure of the model is laid out in Table 4.1. We will discuss these logical building blocks in the following subsections. The model's full program code for Matlab is given in Appendix C.

4.1.2. Forecasting dividends

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.

We assume that all agents make their individual forecasts of dividend dynamics. Dividend flows 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 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 forecasting ability at the systemic level by probabilistic imitation of more successful individuals' behaviour (see Section 4.1.5 for more on this).

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

$$d_{i,y}^{EWMA} = \lambda_1 \cdot d_y + (1-\lambda_1) d_{i,y-1}^{EWMA}. \qquad (4.1)$$

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 (4.1) one time unit would be one year.

Exponential moving averages would clearly be unacceptable estimates of future dividends in a general case. Their function in this model is twofold. First, 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 second, forecasting dividends relative to their moving averages, as opposed to forecasting dividend levels directly, makes forecasting environment more stationary, which facilitates the reinforcement learning task.

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

$$d^{e}_{i,y+n} = d^{EWMA}_{i,y} \cdot a^{div}_{i,y}, \tag{4.2}$$

where $a_{i,y}$ is agent *i*'s dividend expectation 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. The detailed description of the reinforcement learning procedure is provided in Section 3.2. 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.

4.1.3. Estimating fundamental the 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 (*i*) formation of initial estimates of the fundamental value, based on discounted dividend flows, and *(ii)* 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

$$v_{i,y}^{fund} = d_y + \frac{d_{i,y+1}^e}{1+\bar{r}} + ... + \frac{d_{i,y+n}^e}{(1+\bar{r})^n} + \frac{d_{i,y+n}^e / \bar{r}}{(1+\bar{r})^{n+1}}, \tag{4.3}$$

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 smoothened by calculating the exponentially weighted moving averages:

$$v_{i,y}^{EWMA} = \lambda_2 \cdot v_{i,y}^{fund} + (1-\lambda_2) v_{i,y-1}^{EWMA}. \tag{4.4}$$

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 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 (4.4) by multiplying exponentially smoothed estimates of fundamental value by individual price adjustment factors, $a_{i,t}^p$:

$$v_{i,t}^{reserve} = v_{i,t}^{EWMA} \cdot a_{i,t}^p. \tag{4.5}$$

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).

4.1.4. Making individual trading decisions

Having formed their individual beliefs about the value of the stock price, agents have to make specific portfolio rebalancing decisions. In principle, they weigh 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, it 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, it 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.

In this model we assume that investors do not choose to trade at precise reservation prices because in that case they would miss potentially profitable asset allocation opportunities. It is natural to assume that an investor whose perception of the stock value considerably differs from the average market opinion is likely to take advantage of market liquidity and make an order to trade at a price close the prevailing market price rather than to his own reservation price. But what price would it be? The first obvious step, implemented in the model, is to allow limit orders, i. e. orders to trade the security at a specified *or* better price. Given the complexity of agents' interaction, we proceed in the following, intuitively appealing way: (*i*) we determine the possible price quote grid around the prevailing market price (i. e. determine tick sizes and possible price fluctuation bands), (*ii*) estimate aggregate supply and demand schedules, (*iii*) compute each individual's expected end-of-period wealth for every possible trading price and (*iv*) allow agents to make trading decisions that maximise their expected end-of-period wealth.

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 the successful trade for a given price quote 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:

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

2) 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;

3) 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. Figure 4.1 shows a typical example of estimated probabilities of successfully buying and selling one unit of stock at all possible prices (last period's average price set equal to 25 in this relative pricing grid). This particular example reflects an upward-trending market, in which agents reckon they have higher chances (estimated at around 60%) of selling the stock than buying it (estimated at around 40%) at the last period's average price.

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. First, agent i estimates its expected end-of-period stock holdings (i. e. the number of shares) for each possible price quote j:

$$E(h^1_{i,j,t}) = h^0_{i,t} + E(q_{i,j,t}) \cdot b_{i,t} \quad \text{for all } j. \tag{4.6}$$

Here $E(q_{i,j,t})$ denotes the 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 $b_{i,t}$ takes value of 1 if the agent is willing to buy the stock or –1 if it is willing to sell the stock.

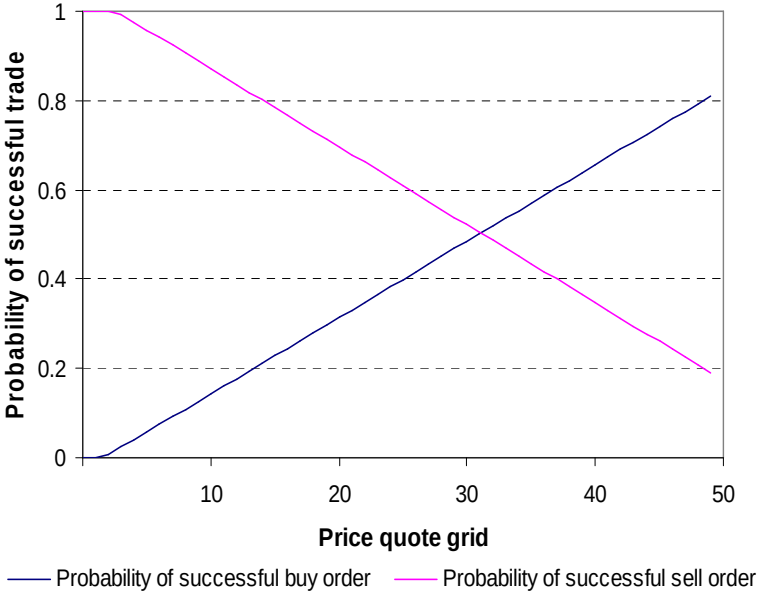


Fig. 4.1. Typical estimated demand and supply schedules in an upward-moving market

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

$$E(m^1_{i,j,t}) = m^0_{i,t} + E(q_{i,j,t}) \cdot x_{j,t} \cdot (-b_i - c) + E(h^1_{i,j,t}) \cdot E(d_{i,t}) \text{ for all } j. \quad (4.7)$$

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 here that the interest on spare cash funds is paid, as well as excess liquidity (cash holdings above some prespecified amount needed for trading) is taken away, at the beginning of the trading period. All of 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 (4.7).

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

$$E(w^1_{i,j,t}) = E(h^1_{i,j,t}) \cdot v^{reserve}_{i,t} + E(m^1_{i,j,t}). \tag{4.8}$$

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:

$$p^q_i = \arg\max_{x_i} E(\vec{w}^1_{i,t}). \tag{4.9}$$

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 trading take place on the centralised stock exchange. The trading mechanism is designed as the double auction system, in which both buyers and sellers submit their competitive orders to trade the stock. More precisely, orders are processed and filled with the help of a simplified order book, which retains a large degree of realism. 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. Traders' strategic interaction within a trading period is hardly important for market dynamics in the much longer term, which is of primary interest in this study.

In this model the order book mechanism works as follows. At the beginning of a trading round, all agents' trade orders are queued randomly and then each of them undergoes a sequential processing procedure. It means that for a given order that is being processed all earlier-queued orders are scanned in search for the most favourable matching (opposite) order. If such an order is found (a tie among several equally good orders is broken arbitrarily), the trade is carried out at the average of the bid and ask price. Otherwise, the order remains open until it makes a match for subsequently processed orders or until the end of the trading period, when it is closed as an unexecuted order. Following the trading round, all agents' cash and securities accounts are updated accordingly.

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. As was mentioned before, it is crucially important for making further trading decisions and it serves as the reference value in the subsequent trading round.

4.1.5. Learning and systemic adaptation in the model

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

We assume that agents' behaviour is driven by the reinforcement learning (more specifically, Q-learning) algorithm since these learning algorithms borrowed from the machine learning literature seem to be conceptually suitable for modelling investor behaviour. 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. The reinforcement-learning agents do just that. 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 if not makes the optimal investment behaviour outright impossible.

The so-called "curse of dimensionality" implies that the straightforward implementation of the basic version of the Q-learning algorithm is rarely possible in complicated environments. Following the standard practice, we apply the Q-learning algorithm with gradient-descent approximation, which is briefly presented in Section 3.2. 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 (4.2)). In each state there are three possible actions – 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 so-called state features, $\vec{\phi}_s$ (see Figure 3.2). 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 basis for successful forecasting. These features include the size of the dividend adjustment factor, relative deviation of the 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:

$$r^{d}_{i,y+n} = -\left(d_{y+n} - E_y(d_{i,y+n})\right)^2. \tag{4.10}$$

Hence, the agent is punished for the 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 the individual stock value estimation, agent *i* also can take one of three actions: fractionally increase or decrease the price adjustment factor, $a^{p}_{i,t}$ (see equation (4.5)), 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, $r^{p}_{i,t+1}$, is simply the log-return on the agent's portfolio:

$$r^{p}_{i,t+1} = \ln\left(h^{1}_{i,t} p_t + m^{1}_{i,t}(1+\bar{r}^{monthly})\right) - \ln\left(h^{0}_{i,t} p_{t-1} + m^{0}_{i,t}\right). \tag{4.11}$$

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 so that they 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. They need not to be risk-averse, as is conventionally assumed about individual consumption-smoothing investors. Indeed, recent developments in extremely turbulent financial markets show that it might well be quite the opposite – in some cases excessive risk-taking might generate superior performance for a prolonged period of time, which in turn generates solid growth in fee income during that time. In addition, it should be noted that in the model an agent's attitude toward risk is determined not only by its reward function but also by evolutionary selection and other systemic adaptation.

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 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.

In this model, dissemination of agents' experience is very stylised. At the end of each period agents are randomly matched in pairs. In every pair, agents' long-term performance measures, which are cumulative past rewards, are compared with each other. If the difference between matched agents' performance measures is sufficiently large (the threshold level is allowed to fluctuate randomly to reflect the random nature of knowledge dissemination), the worse-performing agent simply replicates the more successful agent's experience.

Evolutionary selection is another available option in the present ASM. It assumes bankruptcy of worst-performing agents and their replacement with best-performers. So agents, whose performance relative to the benchmark (which is the average agents' performance) falls below a modeller-specified threshold, go bankrupt. Their place is taken over by best-performers, which then are forced to "split" so that the number of agents remains constant. This has a natural interpretation: inferior fund managers are forced out of the market as unsatisfied clients bring their wealth over to best-performing funds and the latter then have to split for regulatory or any other reasons. Successful agents are given substantial extra rewards in the event of the split, to encourage their good performance.

Finally, the model allows for noise trading behaviour. Unlike in the evolutionary selection, the worst-performers are not replaced by most successful agents. Rather, they scrap their prior learning experience and, as a result, start learning from scratch.

4.2. Simulation results

Like the vast majority of other ASM models, the current model 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 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, model results can offer qualitative insights about market self-regulation, efficiency and other aspects of market functioning. In this section we examine these issues in more detail and report some of the more 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 (presented in Table A3 in Appendix A). The simulation concentrates on altering features of the reinforcement learning, interaction among agents and dividend-generating processes in an attempt to understand importance of intelligent individual behaviour, market setting and population-level changes for the aggregate market behaviour. Other model parameters are kept unchanged. Their values are provided in Table A1.

Dividends are assumed to fluctuate around an exponential trend and their volatility is proportional to the dividend level. The role of the trend is to necessitate the intelligent adjustment of dividend estimates because forecasts based on exponentially weighted moving averages would be clearly biased in this case. Large dividend growth rates can only be sustained over relatively short time horizons, and hence in our very long-term model we have to choose very low dividend growth rates (e. g. 0.15% per year). We also examine deterministic

constant dividends, as a special case (see exact specifications of dividend generating processes in Table A2).

The primary question addressed in most ASM models is the market efficiency issue. Here efficiency is loosely interpreted as the congruence between the stock market price and its fundamentals. In the current setting it is not possible to know the right theoretical stock price, so we basically want to compare the market stock price with risk-neutral estimates of fundamentals.

Let us start with the examination of agents' ability to forecast dividends. Since dividends are driven by very simple data-generating processes, it is not surprising that in the model version with enabled both reinforcement learning and evolutionary selection (Experiment 1 in Table A3) agents are able to form very precise dividend 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 (Experiment 3). 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 prespecified boundaries but perform much worse, taken on 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 dynamics of the market price in relation to the fundamentals. In Experiment 1 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% (see Table A3). Nevertheless, the valuation errors are clearly autocorrelated – 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 relatively even wealth distribution among agents, and each trading period active agents (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 constitutes on average 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 (Experiment 2), 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 remaining 20–30% agents. There are some possible explanations to this overvaluation and wealth concentration. Such overvaluation can be to some 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 worse performing agents 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 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 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 (4.3). 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 no-learning case, 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 the market price fluctuates around them as a result.

The above results confirm 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. The 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 such strong inertia 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.

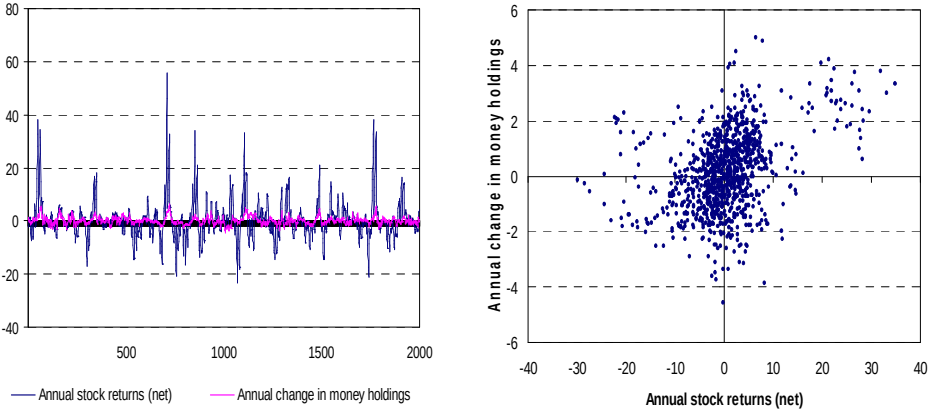


Fig. 4.2. Typical dynamics of percentage stock returns and percentage changes in liquidity in a constant dividend case

Last but not least, we want to examine the relationship between the market price fluctuations and the financial market liquidity. This experiment (see Experiment 4 in Appendix A) 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 excessive price volatility may be the positive relationship between market liquidity and the stock price. Since unnecessary liquidity at the individual level is removed from the system, overall liquidity fluctuates in a haphazard way. Increases in market liquidity bolster solvent demand for the stock and lifts its price. As can be seen from Figure 4.2, 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.

It should be noted that the latter experiment is devised so as to ensure that positive relationship between stock returns (with dividends included) and investors' cash holdings is not linked to fluctuations in dividend payouts, as they

are assumed constant. 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.

4.3. Concluding summary of Chapter 4

In this chapter 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 some 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.

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 learning and the population level adaptation (evolutionary selection in particular) are essential for ensuring any efficiency of the market. The institutional setting alone, such as the centralised exchange based on the double auction trading, cannot ensure effective market functioning. However, market self-regulation ability is found to be weak. 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 find a positive relationship between stock returns and changes in liquidity – there are indications that exogenous shocks to investors' cash holdings lead to strong changes in the market price of the stock.

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 some other ASM models, which are most often based on evolutionary selection procedures and are sometimes criticised for the lack of economic fundament. On the other hand, the proposed model has some weaknesses, which became apparent after fully developing and testing the

model. One of the caveats is that the stock market dynamics strongly depends on perceived risk-neutral fundamentals. Coupled with modeller-specified restrictions, this results in not sufficiently rich market behaviour. Moreover, the economic decision making processes in the model are quite tedious, and some complications appear to contribute little to model's economically interesting results but do constitute an obstacle to model's empirical calibration and testing.

Nevertheless, the proposed modelling approach serves as a basis for a refined and suitable for empirical analysis version of the model, which we develop in the following chapter. The proposed modelling principles could also be expanded and applied for modelling of other markets, such as markets for goods or labour. More generally, similar modelling principles based on complex interaction of adaptive heterogeneous agents will likely form the basis of applied dynamic macroeconomic models.

5

Artificial stock market model version suitable for empirical analysis

In this chapter we present a refined version of the ASM model spelled out in Chapter 4. This version 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 famous 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, 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. We do pursue sensible agent behaviour – by simple means in the uncertain environment. At this stage, theoretically rigorous models are hardly possible in the face of immense complexity of the real world financial markets and limited possibilities to realistically simulate individual human behaviour (and learning). 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 does not rely on any strong assumptions about the formation of aggregate supply or demand (in contrast, in many other ASM models demand for stocks is determined from some theoretical models).

In this chapter we first provide a detailed description of model's main building blocks and simulated market processes. We then conduct simulation experiments in the stationary external environment with the aim of assessing market self-regulation abilities. Further, we apply the model to the actual data and conduct qualitative and quantitative comparison of simulated market dynamics to the actual market dynamics and to the stylised properties of financial markets.

5.1. Description of the model

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.

First of all, we highlight this model's main differences from the model developed in Chapter 4:

1) Forecasting dividends (corporate earnings). In the initial version of the model, dividend forecasts are determined by reinforcement-learning adjustment of past averaged realisations. In the reworked version agents try to identify corporate earnings trends by employing simple econometric tools. This modification enables the analysis of the impact of the historical information filtering on the bubble formation.

2) Estimating risk-neutral fundamentals. One difference is that in the initial version estimation is based on the dividend flow, whereas in this version it is based on expected corporate earnings.

3) Determining individual reservation prices. Initially, reservation prices were based on adjusted fundamental valuations, whereas in the refined

version larger role assumed for the prevailing market consensus (prevailing market price).

4) Reinforcement learning and systemic adaptation. In the current version of the model the learning restrictions are removed (in the initial model, for example, estimates of individual reservation prices were not allowed to deviate too much from the perceived fundamentals), thereby reducing influence of subjective judgement on model results. Evolutionary selection is disabled in the current version.

5) Determining quoted prices in the limit orders. The reworked version of the model is dramatically simplified in this respect, as the quoted prices in limit orders are simply set equal to individual reservation prices.

6) Stock holdings. In the current version of the model, decision-making and trading processes are further simplified by an addition assumption that agents can hold at most one share.

A number of smaller model modifications will also become apparent in the course of describing the current model. Due to the large number of differences with the initial version of the model it is instructive, for the ease of exposition, to present the current model as a new stand-alone model rather than a mere extension or modification of the original model. The remaining similarities and the same modelling principles clearly show that the current version is rooted in the conceptual framework spelled out in the previous chapter.

The model's full program code in Matlab is given in Appendix D.

5.1.1. Model's main building blocks

We start the detailed presentation of the model with a cursory description of the basic market structure, agents and general market processes. The stock market is again 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.

Table 5.1. Main building blocks of the empirical ASM model

Forming individual forecasts of corporate earnings Based on: • Econometric error-correction models of trend-reversion • Individual sample size adjustment via evolutionary selection and pair-wise interaction
Estimating fundamentals and reservation prices Based on: • Discounted expected future corporate earnings (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 Based on: • 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

Model's main processes are summarised in Table 5.1. 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.

5.1.2. 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 on the daily basis 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 of the risky asset 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.

5.1.3. Forming individual forecasts of corporate earnings

Like in the initial version of the model, we hold the view that investors should care about risk-neutral fundamentals of the stock. In order to determine them all agents form their individual forecasts of future corporate earnings, calculate risk-neutral fundamentals and adjust them to obtain their reservation prices.

We assume that 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 non-stationary 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, we want to keep the forecasting procedure as simple as possible. For these reasons agents assume that aggregate earnings follow simple linear trends (that could be related, for example, to the technological progress or 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 regression

$$y_q = \beta_{1,i}^{LR} + \beta_{2,i}^{LR} \cdot q + \varepsilon_{i,q}^{LR} \tag{5.1}$$

and the short-run regression

$$\Delta y_q = \beta_{1,i}^{SR} \Delta y_{q-1} + \beta_{2,i}^{SR} \varepsilon_{i,q-1}^{LR} + \varepsilon_{i,q}^{SR}. \tag{5.2}$$

Here y_q denotes net corporate earnings in quarter q. Betas and epsilons 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 differing 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[11]. Imitation learning at systemic level may occur as individuals observe and copy other people's actions, which 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 and Feltovich, 1999), an individual with highest known utility (Kirchkamp, 2000), etc.

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 become available, agents rerun their regression equations, and in order to do this 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[12] (based on their beliefs about fundamentals)

[11] In the psychology literature it is also known as observational learning (Brenner, 2006).
[12] See how agent performance is defined in Section 1.5.

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 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 and have bet on the market rise. Once market trends change, bearish 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 (5.1) and (5.2). 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. ten 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 zero.

5.1.4. 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 (see equation (4.3) for comparison):

$$v_{i,q}^{fund} = \frac{y_{i,q+1}^{e}}{1+r_{q,q+1}^{f}} + ... + \frac{y_{i,q+n}^{e}}{(1+r_{q,q+n}^{f})^{n}}. \qquad (5.3)$$

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.

Like in the initial version of the model, the risk-neutral 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.

As before, 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 and Froot,1986). 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.

Speaking 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.

We provided this extended discussion of agents' attitude to risk because in the current version of the model the individual reservations price are more loosely tied to perceived risk-neutral fundamentals whilst risk perception at the population level is given a larger role. In this case, agent *i*'s reservation price $v_{i,t}^{reserve}$ in trading period *t* is assumed to be determined by three elements – a change in perceived fundamentals, prevailing market price and the reinforcement learning adjustment factor (compare to equation (4.5)):

$$v_{i,t}^{reserve} = \frac{v_{i,t}^{fund}}{v_{i,t-1}^{fund}} \cdot p_{t-1} \cdot (1 + a_{i,t}). \tag{5.4}$$

The logic behind this pricing equation is simple. Each agent has its individual perceptions about changes in perceived risk-neutral 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 individual learning, which occurs in the process of agent's interaction with environment. Note that in this setting the individual attitude to risk explicitly 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.

5.1.5. Agents' reinforcement learning

We again employ the gradient descent Q-learning algorithm to govern agents' learning and adaptation. Recall this algorithm's principal back-up rule from Chapter 3 (equation 3.15):

$$\vec{\theta}_a \leftarrow \vec{\theta}_a + \alpha[r_{t+1} + \gamma \max_a Q_t(s_{t+1}, a) - Q_t(s_t, a_t)]\vec{\phi}_s, \quad \text{for all } a. \tag{5.5}$$

This back-up rule plays the pivotal role in the actual algorithm determining agents' adaptive behaviour in the present model. As was already mentioned, we cannot expect the convergence guarantee to hold in the complex multi-agent environment. Nevertheless from the economist's viewpoint such behavioural principles are interesting in their own right as they constitute a way to describe sensible behaviour in the highly uncertain environment. In the context of this model, each agent (*i*) adjusts its attitude to risk (takes action), (*ii*) this results in trade orders, then (*iii*) trade outcomes are determined by the entirety of individual trading decisions, (*iv*) this leads to the specific financial outcome (the immediate reward), from which (*v*) some signal about appropriateness of the action can be inferred with the help of back-up rule (5.5).

We can now describe the specific variables that are plugged into this general back-up rule. At time *t* agent *i* takes action $a_{i,t}$ to incrementally adjust its individual reservation price in equation (5.5). 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 the 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:

$$\begin{aligned} m_{i,t}^1 &= (m_{i,t}^0 - b_{i,t} \cdot p_{i,t}) \cdot (1 + r_{t,t+1}^f) + h_{i,t}^1 \cdot d_t, \\ w_{i,t}^1 &= h_{i,t}^1 \cdot p_t + m_{i,t}^1, \\ r_{i,t}^{return} &= w_{i,t}^1 / w_{i,t-1}^1 - 1. \end{aligned} \tag{5.6}$$

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*. 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 (say, 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 e. g. Martinez-Jaramillo, 2007, for discussion). It is natural that professional asset managers shun worse-than-average performance.

5.1.6. Carrying out trades via the centralised exchange

Each trading round agents 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. The decision rule is similar to that of the initial version of the model and is as follows:

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

If $v_{i,t}^{reserve} < p_{t-1}$ *and* $h_{i,t}^{0} = 1 \rightarrow$ *make order to sell 1 share,*

Otherwise, make no order.

The comparison of the reservation price and the market price determines agent's intentions to engage in a trade but it does not pin down the exact price which the agent would choose to submit in a trade order. It is not a trivial problem because less favourable (for the agent) price should normally be associated with higher chance of successful execution of the order (see Section 4.1.4 for discussion). However, in order to avoid complication of the model, we assume very simple agents' behaviour in this respect. 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.

Market price determination and trading take place on the centralised stock exchange through the double auction system, which is identical to that described in Section 4.1.4. 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.

5.2. Simulation in stationary environment and examination of market self-regulation ability

In this subsection 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. It is not a trivial exercise because agents do not know the model of the environment – they only have some, incomplete and subjective, knowledge of economic fundamentals. Market price dynamics only partially depends on (perceived) changes in fundamentals. It is also directly influenced by the reinforcement-learning adjustment, which is aimed at achieving good long-term performance and taking into account behaviour of others. 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 B1 and B2 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.

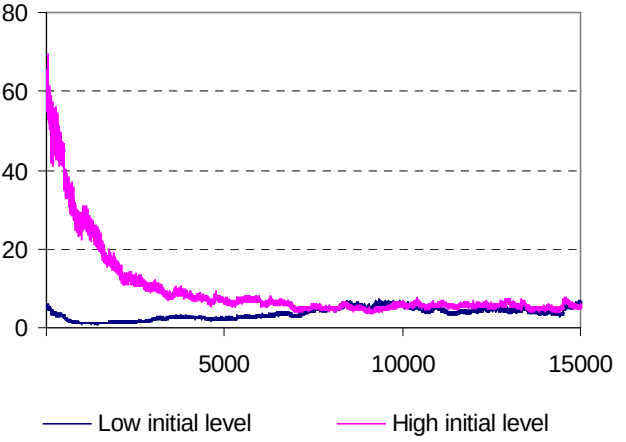


Fig. 5.1. Simulated market price dynamics from different initial price levels (stationary external environment; penalty rate = –1; no loss aversion)

It can be seen from Figure 5.1 that typical 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 an indication of 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 these 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.

Stock prices also approach zero if initial price level is chosen to be very low. Such self-organisation of the system is arguably related to inherent indeterminacy of stock prices. Note that for exogenously given dividend payouts, stock returns are high if either the stock price is low (which gives larger dividend rate) or the stock price is growing (which implies larger capital gains). Hence, low initial price levels and self-reinforcing downward price trends at the beginning of the learning process may bring the system to this equilibrium of the very low stock price and relatively large dividend rates. This serves as an illustration of quite serious difficulties of choosing the appropriate reinforcement-learning reward structure. Tedious experimentation with the specification of reinforcement signals showed that the system often goes the path of "least resistance", which gives logical but economically not very interesting results.

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. In this particular case the simulated stock price is about 3-4 times larger than annual dividends, which implies agents' large degree of risk aversion.

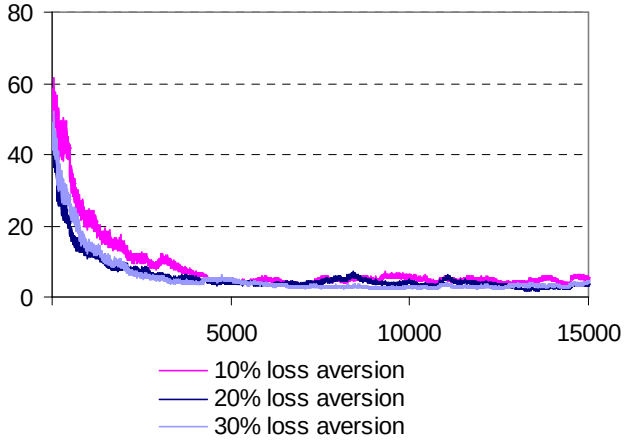


Fig. 5.2. Simulated market dynamics for different loss aversion parameters (stationary external environment; penalty rate = – 2)

Rewarded risk taking can be also supported 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 5.2, 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% – the stock price fluctuates around 4.7, a 20% increase gives the equilibrium price level of roughly 3.7, and 30% asymmetry leads to price fluctuating around 3.1 in this setting.

5.3. Applying the model to the actual data

In this subsection 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 the actual data. Important parameters describing the model and the experimental setup are given in Tables B1 and B3.

5.3.1. Model data and background discussion

In this model the stock can be interpreted as a broad market index, which necessitates having specific empirical data, such as aggregate earnings 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.

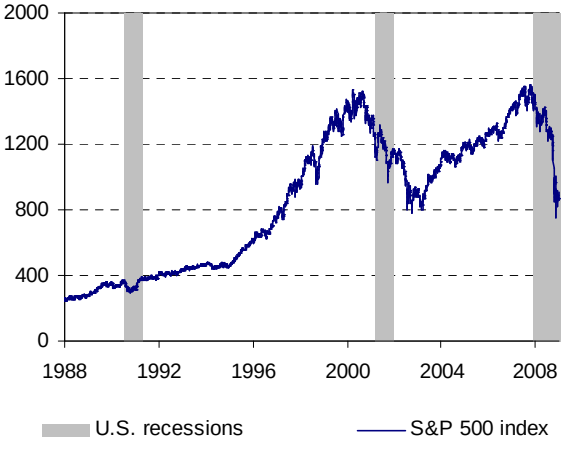


Fig. 5.3. Dynamics of S&P 500 index and official timing of U.S. recessions

Source: Standard and Poor's, National bureau of Economic Research.

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 and 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 5.3).

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 1565 on October 9, 2007. Something that started with stagnation of the U.S. housing market, failures of highly leveraged market players 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% in one week starting October 6. From peak to trough (recorded on March 9, 2009) the S&P 500 lost 57%. 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.

Let us examine fundamentals of the S&P 500 index a bit more closely. As can be seen from Figure 5.4, there was indeed an upward shift in the earnings trend around the time when the first stock market boom began. The trend ended abruptly with the burst of the dot-com bubble and the onset of the recession. During the second boom episode earnings exhibited even brisker growth. Again

in line with the timing of the stock market boom, they declined dramatically in the second half of 2007 and later, after some rebound, plunged deep into the negative territory.

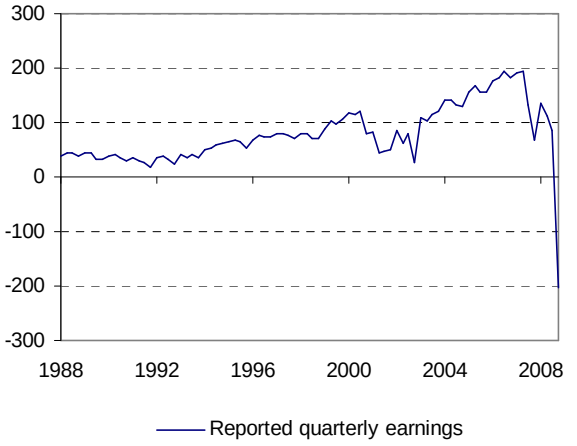


Fig. 5.4. Quarterly corporate earnings of companies included in S&P 500 index (bn U.S. dollars)

Source: Standard and Poor's.

For a more comprehensive assessment of fundamentals it is necessary to take into account the interest rate environment, as foregone interest rates constitute alternative costs to the risky investment[13]. For this purpose we calculate the inverted price-to-earnings ratios (with earnings summed over the last four quarters) and plot them against the one-year Treasury yields in Figure 5.5. Note that as stockholders are entitled to company's earnings, the inverted price-to-earnings ratio reflects the fundamental part of the return on their stock holdings and is directly comparable to returns on alternative investments. One obvious observation from Figure 5.5 is that the earnings indicator closely tracks the one-year Treasury yield. In this sample the difference between the two indicators on average equals zero and the correlation coefficient between the two is 0.66. This empirical observation does not fit in well with the neoclassical paradigm of rewarded risk.

[13] Recall that this basic idea is also very important in our proposed model.

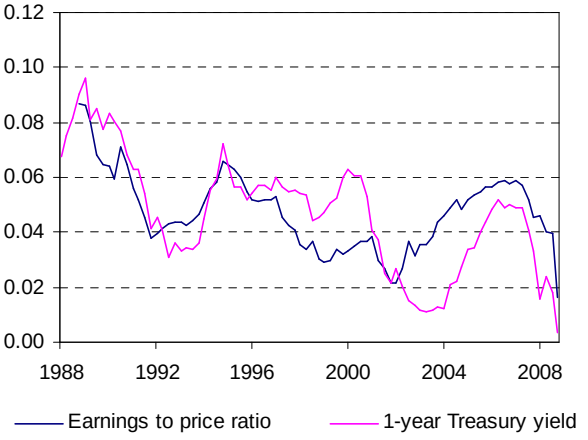


Fig. 5.5. Corporate earnings as a fraction of market value compared to riskfree returns

Source: U.S. Federal Reserve, Standard and Poor's.

It is also interesting to note that periods of larger discrepancy between these indicators coincide with boom episodes. Importantly, the gaps of the opposite signs suggest that the two boom episodes were of different nature. During the technology boom the abovementioned measure of stock returns was considerably lower than the riskfree rate, indicating that actual corporate earnings could not justify relatively high stock prices. Instead, the market beliefs about the immense impact of the technological breakthrough on future corporate earnings pushed the prices even higher, which compensated insufficient corporate earnings with capital gains. In contrast, during the latter market boom the indicator of fundamental stock returns was consistently higher than the riskfree rate owing to both improved corporate earnings and the extremely low interest rates (determined by the monetary policy stance and low inflation environment). The large positive gap made investing in stocks very attractive, and the boom was further stimulated by self-reinforcing stock price rises. In both cases the bubbles burst with the collapse of company earnings, which proved that earnings expectations were unjustified. Also, in the latter episode the monetary policy tightening may have been another major determinant of the market turnaround.

5.3.2. 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 5.6 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.

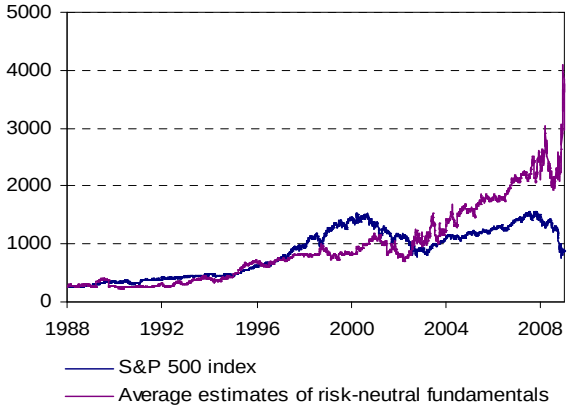


Fig. 5.6. Average simulated estimate of fundamental value compared to S&P 500 index

Source: Standard and Poor's and 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 5.7 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.

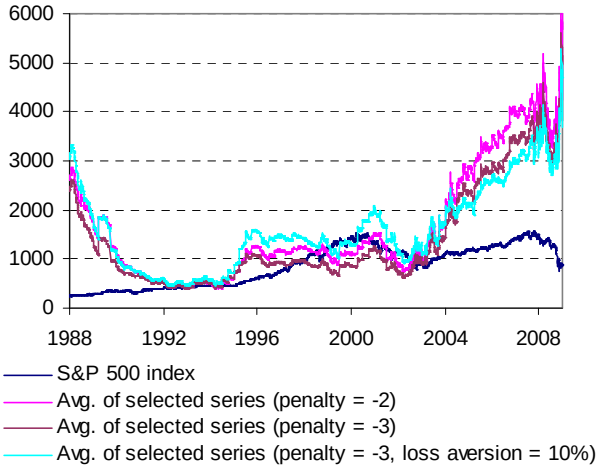


Fig. 5.7. Simulated market prices compared to S&P 500 index

Source: Standard and Poor's and 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.

5.3.3. 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 1000 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 5.8) 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 B3 in the Appendix). Simulated daily price fluctuations can reach up to 40%, whereas maximum actual fluctuations were about 10% 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% to 5%. 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% in

absolute value from the sample as outliers (they constitute about 0.5% of the effective sample).

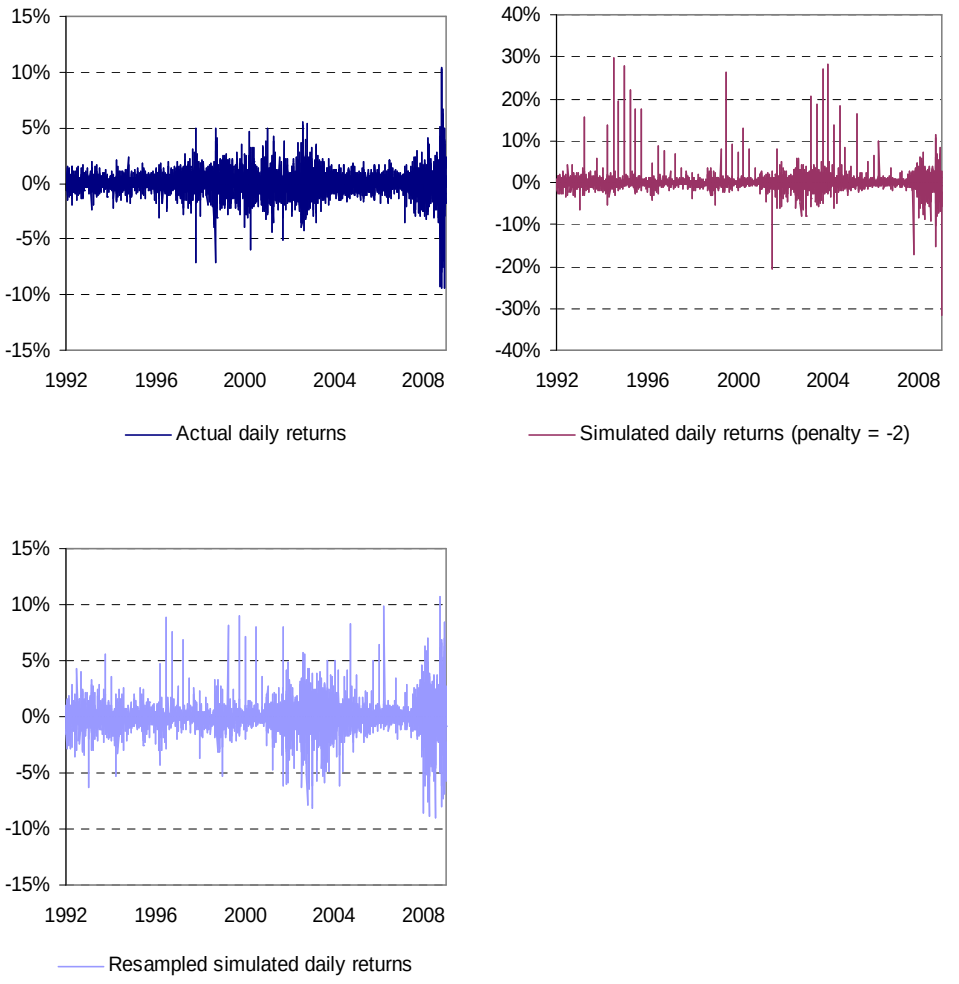


Fig 5.8. Simulated and actual daily returns

Source: Standard and Poor's and 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 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 B3). 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.

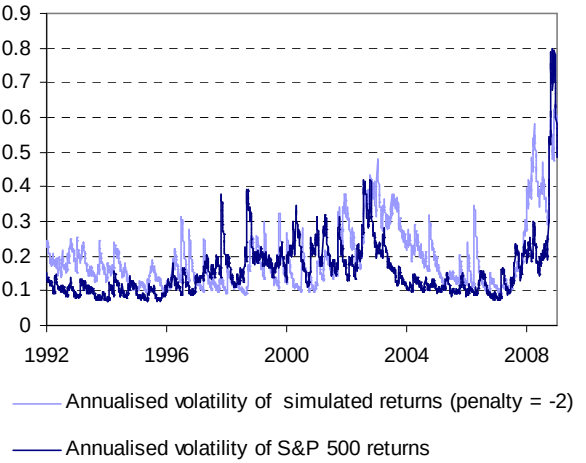


Fig. 5.9. Simulated and actual annualised conditional volatility of stock returns

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

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 5.9). 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.

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 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.

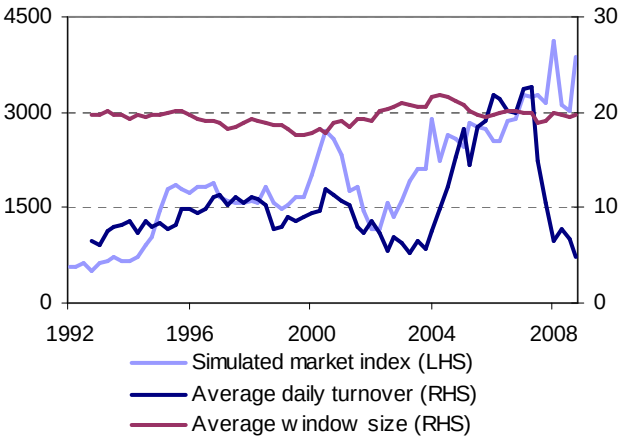


Fig. 5.10. Simulated market liquidity (no. of transactions) and window or relevant history (in quarters) compared to market price level

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 5.10). 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.

5.4. Concluding summary of Chapter 5

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 did 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 (see, e. g., Arthur, 1995 for discussion). 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. More specifically, agents individually determine corporate earnings trends, form risk-neutral valuations of the stock and adjust these estimates by taking into account perceived changes in fundamentals, general market sentiment and their own performance. Moreover, there are no strong assumptions about supply and demand formation in the model environment, and they are determined purely by the interaction of agents. Other realistic features of the model include daily trading and managing of accounts, quarterly dividends, centralised stock exchange based on competitive limit orders, etc.

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.

General conclusions

Concluding summary and comments

Global economic and financial problems caused by the disorderly unwinding of imbalances that had been accumulating over decades cannot be easily reconciled with the efficient market paradigm based on rational expectations and perfectly rational representative agent assumptions. The collapse of global consumption can even be seen as a result of the disregard of intertemporal budget constraints, which is of course at odds with basic principles of consumer (investor) optimising behaviour. In any case, the mainstream financial and economic theory cannot explain current economic developments, give adequate forecasts or policy prescriptions. In this context there is a growing need to look for a replacement for the work-horse financial and macroeconomic models based on heroic assumptions.

This work advocates the view that agent-based financial modelling, which takes financial markets as complex dynamical systems consisting of interacting heterogeneous agents, can potentially become a viable alternative. However, significant further interdisciplinary progress is needed before agent-based models can take the centre-stage of financial market modelling.

We also argue that it is necessary to take into account market complexity, agent heterogeneity, bounded rationality and adaptive (though not simplistic) expectations. These features are usually absent in standard financial models but are at the heart of agent-based modelling. Agent-based models use the procedural (computer code) – rather than mathematical – model description form, which technically enables researchers to easily inquire into many interesting features of systemic behaviour. However, theoretical or empirical foundation of individual agents' behaviour and external validation of model results are two main problem areas, which have not been systemically addressed to date and constitute a serious obstacle to the further progress of agent-based financial modelling.

In the brief review of the literature on the artificial stock market modelling two broad categories of models are presented: (i) models based on stochastic, heuristic and standard theory-implied behavioural rules and (ii) models with learning agents or evolutionary systemic adaptation. The first group of models generally emphasise the role of agent heterogeneity in determining complex market behaviour and emergent systemic properties, whereas the second group of models concentrate on the generation of good investment strategies and the macro-level implications of intelligently adapted investment strategies.

To our knowledge, no full-fledged artificial stock market models have been developed by Lithuanian researchers. However, in this work we provide some references to somewhat related studies conducted by Lithuanian scholars. They mostly fall into one of these areas: research of investment decision making, analysis of emergent properties of stock markets, and analysis of empirical relationship between market dynamics and macroeconomic variables.

Intelligent adaptation in the highly uncertain environment can arguably be key to understanding actual financial market behaviour. One of the possibilities to model agent behaviour is to resort to the artificial intelligence literature for specific algorithms of adaptation and learning. In particular, we find reinforcement learning algorithms economically very appealing, even though there are certain problems with their practical implementation. Without knowing the true model of reality, reinforcement-learning agents learn from interaction with environment and adjust their strategies so that they attain maximum long-term reinforcement (utility) from the environment. Similarly, investors seek to reach their long-term objectives in highly uncertain environment.

In our ASM modelling approach we emphasise the importance of the economic content of agent-based models and reckon that overly simplistic or "black-box" modelling do not greatly enhance generative understanding of financial market processes. Though agents' behaviour in our proposed models is still largely ad hoc like in most ASM models, we put a lot of effort to ensure that

their basic behavioural principles accord to the commonsensical understanding of sensible investment behaviour in the highly uncertain environment.

In this work we developed two closely related ASM models. The initial version of the model is more detailed and largely served as an organising framework for a more refined analysis. The more parsimonious and simpler version was partly calibrated to empirical data and was used to enhance generative understanding of two famous bubble episodes during the last two decades.

The basic ASM model, spelled out in Chapter 4, is based on the interaction of heterogeneous agents whose forward-looking behaviour is driven by the reinforcement learning algorithm combined with some 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.

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.

One of the main problems with this model is that the stock market dynamics strongly depends on perceived risk-neutral fundamentals. Coupled with modeller-specified restrictions, this results in not sufficiently rich market behaviour. The investment decision-making processes in the model are quite tedious, and some complications appear to contribute little to model's economically interesting results but do constitute an obstacle to model's empirical calibration and testing. Nevertheless, the proposed modelling approach serves as a basis for a refined and suitable for empirical analysis version of the model.

In the second, refined model agents also exhibit inductive behaviour, as they form relatively simple models of the world, act according to them and update those models in order to reach their goals. More specifically, agents individually determine corporate earnings trends, form risk-neutral valuations of the stock and adjust these estimates by taking into account perceived changes in fundamentals, general market sentiment and their own performance. This adjustment is again governed by the reinforcement learning algorithm. Just like in the original version of the model, there are no strong assumptions about supply and demand formation in the model environment, and they are determined purely by the interaction of agents. Other realistic features of the

model include daily trading and managing of accounts, quarterly dividends, centralised stock exchange based on competitive limit orders, etc.

We conducted twofold simulations with the model. First, we examined some properties of the simulated market in a controlled stationary environment. The model was also (partly) calibrated to the U.S. financial data, and properties of simulated market dynamics were compared to those of the S&P 500 index.

Overall, model results are 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 but it does promise to become one of the main alternatives to the crisis-laden neoclassical financial theory.

Main conclusions

1. An original ASM model is proposed. Based on it, a parsimonious partially calibrated witg empirical data model version is developed. The proposed models provide a basic framework for the generative analysis of financial market processes and can serve as a simulation tool of market processes. The proposed agent-based modelling framework can be further modified for use in practical portfolio management applications.
2. Previously little explored possibilities to employ reinforcement learning algorithms (more specifically, Q-learning) in ASM models are implemented. In this work we provide arguments for the conceptual relevance of these behavioural algorithms in the domain of economic problems.
3. Emergent properties of the proposed artificial stock markets are in discord with the neoclassical financial theory's postulates on informational efficiency and homogeneous agent behaviour.
4. The market price of the stock in the basic version of the model broadly in line with fundamentals but over- or under-valuation runs are sustained for prolonged periods.
5. In the basic version of the model both individual learning and adaptation at the population level (more specifically, evolutionary selection) are essential for ensuring any efficiency of the market. The institutional setting alone, such as the centralised exchange based on the double auction trading, cannot ensure effective functioning of the market.

6. Market self-regulation ability in the basic version of the model is found to be weak. Even in the case of activated 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' subjective beliefs for the market price dynamics.

7. Experiments with the basic model confirm a positive relationship between stock returns and changes in liquidity – there are indications that exogenous shocks to investors' cash holdings lead to strong changes in the market price of the stock.

8. Experiments with the refined model in the controlled environment confirm market's self-regulation abilities, as different initial market prices tend to converge to similar levels associated with the same underlying fundamentals.

9. 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.

10. In the case when the model is calibrated to the U.S. financial data, 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 than the actual breaks.

11. 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 real interest rate environment played a larger role.

12. The proposed ASM models demonstrate several plausible mechanisms of financial bubble formation: they include excess liquidity, earnings forecast errors that get dispersed owing to specific agent communication ("herd instincts") and competitive pressures to concentrate on short-term gains against the backdrop of bubble episodes ("trend riding").

References

Arthur, W. B. 1995. Complexity in Economic and Financial Markets, *Complexity* 1(1): 20–25.

Arthur, W. B.; Holland, J.; LeBaron, B.; Tayler, P. 1997. Asset Pricing under Endogenous Expectations in an Artificial Stock Market, in: *The Economy as an Evolving Complex System II*. Eds. W. B. Arthur, S. Durlauf, D. Lane, Reading, MA: Addison-Wesley, 15–44.

Axelrod, R.; Tesfatsion, L. 2006. A Guide for Newcomers to Agent-based Modelling in the Social Sciences, in *Handbook of Computational Economics 2*. Eds. L. Tesfatsion, K. L. Judd. Amsterdam: Elsevier, 1648–1659.

Beltratti, A.; Margarita, S. 1993. Evolution of Trading Strategies among Heterogeneous Artificial Economic Agents, in *Proceedings of the Second International Conference From Animals to Animats 2*. Honolulu, 1993, 494-501.

Bergh, van den W.M.; Boer, K.; de Bruin A.; Kaymak, U.; Spronk, J. 2002. On Intelligent Agent-Based Analysis of Financial Markets, Rotterdam: Erasmus University Working Paper.

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

Bowling, H. M.; Veloso, M. M. 2001. Convergence of Gradient Dynamics with a Variable Learning Rate, in *Proceedings of the Eighteenth International Conference on Machine Learning*, Williamstown, MA, 27–34.

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

Campbell, J. Y.; Shiller, R. J. 1988. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1: 195–1227.

Chen, S. H.; Yeh C. H., 2002a. Evolving Traders and the Business School with Genetic Programming: A New Architecture of the Agent-Based Artificial Stock Market, *Journal of Economic Dynamics and Control* 25: 363–394.

Chen, S. H.; Yeh, C. H. 2002b. On the Emergent Properties of Artificial Stock Markets: the Efficient Market Hypothesis and the Rational Expectations Hypothesis, *Journal of Economic Behavior and Organization* 49: 217–239.

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

Cutler, D. M;, Poterba J. M.; Summers L. H. 1989. What Moves Stock Prices? *Journal of Portfolio Management* 15: 4–12.

De Grauwe, P.; Dewachter, H.; Embrechts, M. 1993. *Exchange Rate Theory: Chaotic Models of Foreign Exchange Markets*. Oxford: Blackwell.

DeLong, J. B.; Shleifer, A.; Summers, L. H.; Waldmann, R. J. 1990a. Noise Trader Risk in Financial Markets, *Journal of Political Economy* 98: 703–738.

DeLong, J. B.; Shleifer, A.; Summers, L. H.; Waldmann, R. J. 1990b. Positive Feedback Investment Strategies and Destabilizing Rational Speculation, *Journal of Finance* 45: 379–395.

Dubauskas, G.; Teresienė, D. 2005. Autoregressive Conditional Skewness, Kurtosis and Jarque-Bera in Lithuanian Stock Market Measurement, *Engineering Economics* 5 (45): 19–24.

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.

Duffy, J. 2006. Agent-based Models and Human Subject Experiments, in *Handbook of Computational Economics 2*. Eds. L. Tesfatsion, K. L. Judd. Amsterdam: Elsevier, 949–1011.

Epstein, J. M. 1999. Agent-Based Computational Models and Generative Social Science, *Complexity* 4: 51–57.

Epstein, J. M. 2006. Remarks on the Foundations of Agent-Based Generative Social Science, in *Handbook of Computational Economics 2*. Eds. L. Tesfatsion L., K. L. Judd. Amsterdam: Elsevier, 1585–1604.

Eshel, I.; Samuelson, L.; Shaded A. 1998. Altruists, Egoists and Hooligans in Local Interaction Model, *American Economic Review* 88: 157–179.

Farmer, J. D.; Geanakoplos J. 2008. The Virtues and Vices of Equilibrium and the Future of Financial Economics, Cowles Foundation Discussion Paper No. 1647. Available at http://ssrn.com/abstract=1112664 (accessed on 10 March 2009).

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. 1986. Short-Term and Long-Term Expectations of the Yen/Dollar Exchange Rate: Evidence from Survey Data, NBER Working Paper 292. Available at http://www.federalreserve.gov/pubs/ifdp/1986/292/ifdp292.pdf (accessed on 22 March 2009).

Frankel, J. A.; Froot, K. A. 1988. Explaining the Demand for Dollars: International Rates of Return and the Expectations of Chartists and Fundamentalists, *Agriculture, Macroeconomics, and the Exchange Rate*. Eds. R. Chambers, P. Paarlberg. Boulder: Westview Press, 88–102.

Franklin, S.; Graesser, A. 1996. Is it an Agent, or just a Program? A Taxonomy for Autonomous Agents, in *Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages* 3, Berlin: Springer-Verlag, 21–35.

Girdzijauskas, S.; Štreimikienė D.; Čepinskis, J.; Moskaliova V.; Jurkonytė, E; Mackevičius, R. 2009. Formation of Economic Bubbles: Causes and Possible Preventions, *Baltic Journal on Sustainability*, 15 (2): 267–280.

Girdzijauskas, S.; Štreimikienė, D. 2008. Logistic Growth Models for Analysis of Stock Market Bubbles, in *Proceedings of International Conference of Financial Engineering*, London, 1166–1170.

Gode, D. K.; Sunder, S. 1993. Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality, *Journal of Political Economy* 101: 119–137.

Grossman, S.; Stiglitz, J. 1980. On the Impossibility of Informationally Efficient Markets, *American Economic Review* 70: 393–408.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*, Ann Arbor: University of Michigan Press.

Hommes, C. H. 2006: Heterogeneous Agent Models in Economics and Finance, in *Handbook of Computational Economics 2*. Eds. Tesfatsion, L., Judd K. L. Amsterdam: Elsevier, 1109–1186.

Hu, J.; Wellman, M. P. 2003. Nash Q-Learning for General-Sum Stochastic Games, *Journal of Machine Learning Research* 4: 1039–1069.

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.

Jacobs, B. I.; Levy, K. N.; Markowitz, H. M. 2004. Financial Market Simulation in the 21st Century, *The Journal of Portfolio Management* 30: 142–151.

Jermann, U.; Quadrini V. 2002. Stock Market Boom and the Productivity Gains of the 1990s, BNER Working Paper No. W9034. Available at http://www.nber.org/papers/w9034 (accessed on 17 April 2009).

Joshi, S.; Parker, J.; Bedau, M. A. 2000. Technical Trading Creates a Prisoner's Dilemma: Results from an Agent-Based Model, in *Proceedings of Computational Finance*, New York, 465–479.

Kaelbling, L. P.; Littman, M. L.; Moore, A. W. 1996. Reinforcement Learning: A Survey, *A Journal of Artificial Intelligence Research* 4: 237–285.

Kahneman, D.; Tversky, A. 1973. On the Psychology of Prediction, *Psychological Review* 80: 237–251.

Kahneman, D.; Tversky, A. 1979. Prospect Theory: An Analysis of Decision under Risk, *Econometrica* Vol. 47 (2): 263–292.

Keynes, J. M. 1936. *The General Theory of Unemployment, Interest and Money*, New York: Harcourt, Brace and World.

Kim, G.; Markowitz, H. 1989. Investment Rules, Margin, and Market Volatility, *Journal of Portfolio Management* 16: 45–52.

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

Kirman, A. P. 1991. Epidemics of Opinion and Speculative Bubbles in Financial Markets, in *Money and Financial Markets*, Ed. M. Taylor. London: Macmillan, 354–368.

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.

Kvedaras, V.; Basdevant O. 2004. Testing the Efficiency of Emerging Markets: The Case of the Baltic States, *Journal of Probability and Statistical Science* 2: 111–138.

LeBaron, B. 2001a. Evolution and Time Horizons in an Agent-Based Stock Market, *Macroeconomic Dynamics* 5: 225-254.

LeBaron, B. 2001b. A Builder's Guide to Agent-Based Financial Markets, *Quantitative Finance* 1(2): 254–261.

LeBaron, B. 2003. Calibrating an Agent-Based Financial Market to Macroeconomic Time Series, Brandeis University Working paper. Available at http://www.hks.harvard.edu/netgov/files/complexity/Blake%20Paper.pdf (accessed on January 5, 2007)

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

LeBaron, B.; Arthur, W. B.; Palmer, R. 1999. Time Series Properties of an Artificial Stock Market, *Journal of Economic Dynamics and Control* 23: 1487–1516.

Lettau, M. 1997. Explaining the Facts with Adaptive Agents: The Case of Mutual Fund Flows, *Journal of Economic Dynamics and Control* 21: 1117–1148.

Littman, M. 1994: Markov Games as a Framework for Multi-Agent Reinforcement Learning, in *Proceedings of the Eleventh International Conference on Machine Learning*, New Brunswick, NJ, 157–163.

Lo, A. W.; MacKinlay, A. C; 1988. Stock Prices Do not Follow Random Walks: Evidence from a Simple Specification Test, *Review of Financial Studies* 1: 41–66.

Lux, T. 1995. Herd Behaviour, Bubbles and Crashes, *Economic Journal* 105: 881–896.

Lux, T.; Marchesi, M. 1999. Scaling and Criticality in a Stochastic Multi-Agent Model of a Financial Market, *Nature* 397: 498–500.

Lux, T.; Marchesi, M. 2000. Volatility Clustering in Financial Markets: A Micro-Simulation of Interacting Agents, Discussion Paper Series B 437. Bonn: University of Bonn.

Martinez-Jaramillo, S. 2007. Agent Based Approach to Reproduce Stylized Facts and to study the Red Queen Effect, PhD dissertation, University of Essex.

Mehra, R.; Prescott, E. 1985. The Equity Premium: A Puzzle, *Journal of Monetary Economics* 15: 145–161.

Mitchell, T. 1997. *Machine Learning.* McGraw Hill.

Moody, J.; Saffell, M. 2001. Learning to Trade via Direct Reinforcement, *IEEE Transactions on Neural Networks* 12(4): 875–889.

Muth, J. F. 1961. Rational Expectations and the Theory of Price Movements, *Econometrica* 29: 315–335.

Nenortaitė J. 2007. A Particle Swarm Optimization Approach in the Construction of Decision-Making Model, *Information Technology and Control* 36 (1A): 158–163.

Nenortaitė, J.; Simutis, R. 2005. Adapting particle swarm optimization to stock markets, in *Proceedings of Intelligent Systems Design and Applications: 5th International Conference*, Wroclaw, 2005, 520–525.

Neuneier, R. 1996. Optimal Asset Allocation using Adaptive Dynamic Programming, in *Advances in Neural Information Processing System* Eds. D. S. Touretzky, M. C. Mozer, M. E. Hasselmo. Cambridge, MA: MIT Press, 952–958.

Pilinkus, D. 2009. Stock Market and Macroeconomic Variables: Evidences from Lithuania, *Economics and Management* 14: 884–891.

Plikynas, D. 2008. Portfolio Design and Optimization Using Neural Network Based Multiagent System of Investing Agents, in *Proceedings of 20th EURO Mini Conference "Continuous Optimization and Knowledge-Based Technologies"*, Neringa, 2008, 137– 142.

Plikynas, D.; Simanauskas, L.; Būda, S. 2002. Research of Neural Network Methods for Compound Stock Exchange Indices Analysis, *Informatica* 13 (4): 465–484.

Ramanauskas, T. 2007. Interest Rates, Credit and Macroeconomy – a Small Structural Model for the Lithuanian Case. *Conference on Monetary Policy Challenges Resulting from Rapid Credit Growth in Converging Economies*, Warsaw, 2007.

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

Ramanauskas, T. 2009. Empirical Version of an Artificial Stock Market Model, *Monetary Studies* 1: 5–26.

Ramanauskas, T.; Jakaitienė, A. 2007. Namų ūkių vartojimo modeliavimas, *Pinigų studijos* 1: 25–44.

Roth, A. E.; Erev, I. 1995. Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term, *Games and Economic Behavior* 8: 164–212.

Routledge, B. R. 1999. Adaptive Learning in Financial Markets, *Review of Financial Studies* 12: 1165–1202.

Routledge, B. R. 2001. Genetic Algorithm Learning to Choose and Use Information, *Macroeconomic Dynamics* 5: 303–325.

Rutkauskas, A. V. 2008. Investor's Possibilities Evaluation in Capital and Exchange Markets, in *Proceedings of 5th International Scientific Conference "Business and Management"*, Vilnius, 281–288.

Rutkauskas, A. V.; Ramanauskas, T. 2009. Agent-based Stock Market Simulation Model, *Intellectual Economics* 2 (6): 83–95.

Samanidou, E.; Zschischang, E.; Stauffer, D.; Lux, T. 2007. Agent-Based Models of Financial Markets, *Reports on Progress in Physics* 70: 409–450.

Sethi, R.; Franke, R. 1995 Behavioural Heterogeneity under Evolutionary Pressure: Macroeconomic Implications of Costly Optimization, *Economic Journal* 105: 583–600.

Sharpe, W. F. 2007. *Investors and Markets: Portoflio Choices, Asset Prices, and Investment Advice*. Princeton, NJ: Princeton University Press.

Shiller, R. J. 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends, NBER Working Paper 0456. Available at http://www.nber.org/papers/w0456.pdf (accessed on 3 August 2007).

Shiller, R. J. 1984. Stock Prices and Social Dynamics, *Brookings Papers in Economic Activity 2*, 457–510.

Simon, H. A. 1957. *Models of Man*. New York: Wiley.

Simutis R.; Masteika S. 2004. Intelligent stock trading systems using fuzzy-neural networks and evolutionary programming methods, *Solid State Phenomena*, Vol. 97–98: 59–64.

Simutis R.; Masteika, S. 2006. A hybrid stock trading strategy based on fundamental and technical analysis, in *Proceedings of Databases and Information Systems: Seventh International Baltic Conference*, *Vilnius*, 2006, 113–121.

Stankevičius, G. 2001. Forming of the Investment Portfolio Using the Self-organizing Maps (SOM), *Informatica* 12 (4): 573–584.

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

Tay, N. S. P.; Linn, S. C. 2001. Fuzzy Inductive Reasoning, Expectation Formation and the Behavior of Security Prices, *Journal of Economic Dynamics and Control* 25: 321–362.

Teresienė, D.; Aarma, A.; Dubauskas, G. 2008. Relationship between Stock Market and Macroeconomic Volatility, *Transformations in Business & Economics* Vol. 7, No. 2 (14): 102–114.

Tesauro, G. 2002. Multi-Agent Learning Mini-Tutorial, in Workshop on Multi-Agent Learning, Vancouver, 2002. Available at *www.cs.rutgers.edu/~mlittman/topics/nips02/nips02/tesauro.ppt* (accessed on 20 February 2008).

Tesauro, G. 2004. Extending Q-Learning to General Adaptive Multi-Agent Systems, *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 25 – 33.

Tesauro, G.; Kephart, J. O. 2002. Pricing in Agent Economies Using Multi-Agent Q-Learning, *Autonomous Agents and Multi-Agent Systems* 5: 289–304.

Tesfatsion, L. 2006. Agent-based Computational Economics: A Constructive Approach to Economic Theory, in *Handbook of Computational Economics 2*. Eds. L. Tesfatsion, K. L. Judd, 831–893. Amsterdam: Elsevier.

Tversky, A.; Kahneman, D. 1974. Judgement under Uncertainty: Heuristics and Biases, *Science* 185: 1124–1131.

Van Roy, B. 1999. Temporal-Difference Learning and Applications in Finance, in *Computational Finance (Proceeding of the Sixth International Conference on Computational Finance*. Eds. Y. S. Abu-Mostafa; B. LeBaron; A. W. Lo; A. S. Weigend. Cambridge, MA: MIT Press, 121-130.

Watkins, C. J. C. H. 1989. Learning from Delayed Rewards. PhD dissertation. Cambridge, UK: University of Cambridge, King's College.

Author's publications related to the dissertation research

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

Ramanauskas, T. 2009. Empirical Version of an Artificial Stock Market Model, *Monetary Studies* (*Pinigų studijos*), 1: 5–26.

Rutkauskas, A. V.; Ramanauskas, T. 2009. Agent-based Stock Market Simulation Model, *Intellectual Economics* (*Intelektinė ekonomika*), 2 (6): 83–95.

Appendices

The content of appendices is available on the enclosed disc

Appendix A. Parameter setting and experimental results of the initial model

Table A1. Key parameter settings of the ASM model

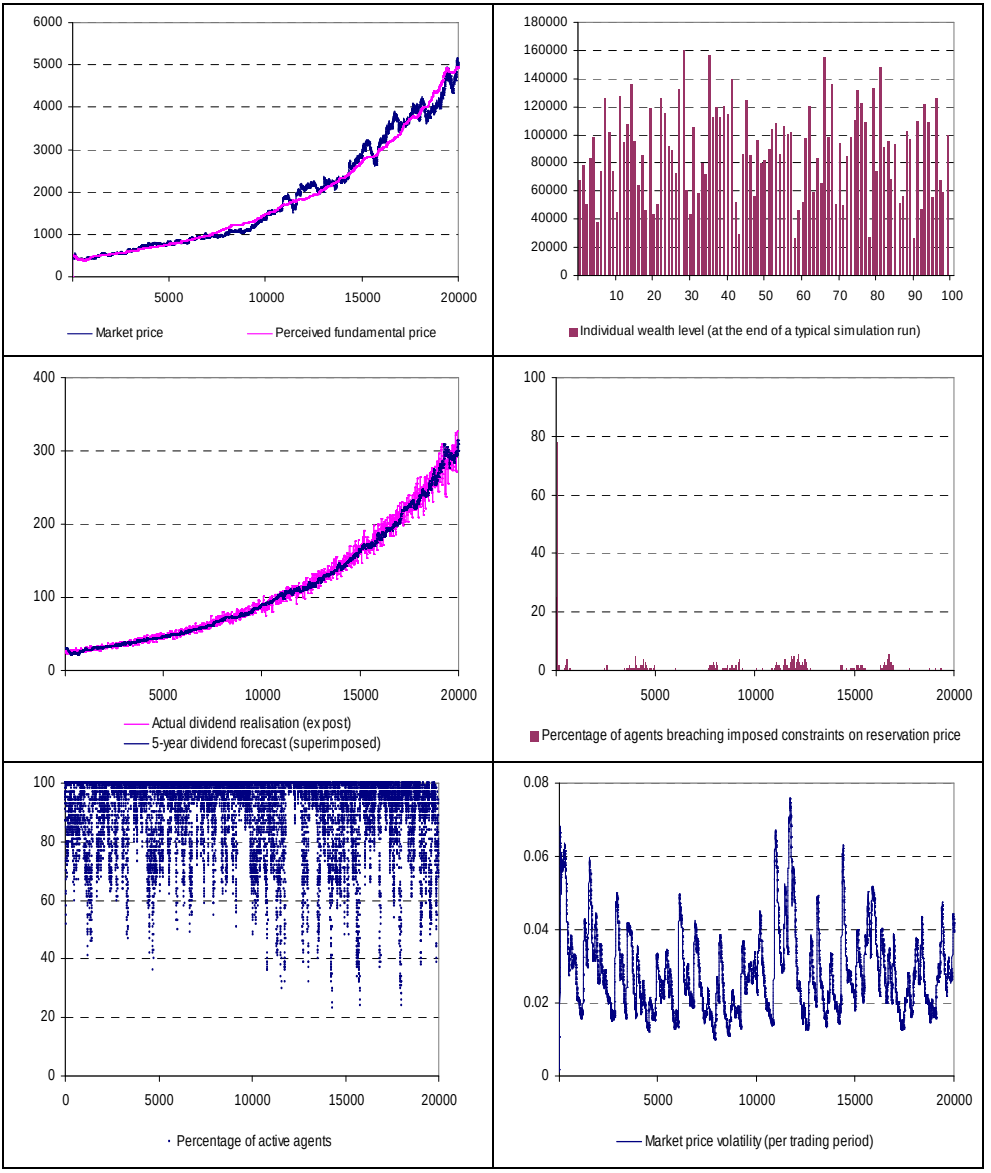
General parameters	
Length of a simulation run (number of trading periods in a run)	20000
Number of simulation runs in a batch	10
Number of agents	100
Total number of shares	10000
Frequency of dividend payouts	Annual
Monthly discount rate	0.995
Annual interest rate on bank account	0.062
Liquidity ceiling (as a multiple of current stock price)	5
Trading	
Number of feasible price quotes in a trading period	50
Frequency of trading rounds	Monthly
Trade cost (as a fraction of trade value)	0.001
Learning	
Learning rate (alpha)	0.1
Exploration rate (epsilon)	0.1
Subjective discount parameter of reinforcement learning	0.995
Dividend forecasting horizon	5 years
Smoothing parameter in the EWMA of dividends, fundamental value	0.1
Dividend forecast constraint (as a fraction of current dividend)	$\pm$ 0.3
Individual reservation price constraint (as a fraction of perceived fundamentals)	$\pm$ 0.2
Action step size in the process of dividend learning (allowed percentage changes of the dividend adjustment factor)	-0.02; 0; 0.02
Action step size in the process of reservation price formation (allowed percentage changes of the price adjustment factor)	-0.02; 0; 0.02
Bankruptcy conditions in evolution procedure (and noise trading)	
Maximum number of bankruptcies in a trading round	3
Performance threshold (as a percentage of average performance)	0.7
Threshold for strategy imitation	
Average difference between two compared strategies (as percentage of the leading strategy)	0.2

Table A2. Specification of model experiment runs

Dividend generating process (Model 1)	$div_t = 25 \cdot 1.000125^t + 0.05 \cdot div_{t-1} \cdot \varepsilon_t$		
Dividend generating process (Model 2)	$div_t = 25$		
	Model	***Learning***	***Evolution***
Experiment 1	Model 1	ON	ON
Experiment 2	Model 1	ON	OFF
Experiment 3	Model 1	OFF	ON
Experiment 4	Model 2	ON	ON

Table A3. Basic descriptive statistics of simulation experiments

Experiment	***1***	***2***	***3***	***4***
Dividend forecasting				
Average forecast bias, %	-0.1	-0.1	-0.8	0.0
Average absolute forecast error, %	0.4	0.4	1.4	0.1
Price dynamics relative to perceived fundamentals				
Average price bias from fundamentals, %	-1.6	5.9	7.6	-0.1
Average length of overvaluation runs	43.7	405.9	63.2	63.5
Average length of undervaluation runs	59.9	2.9	2.8	62.4
Upper semi-deviation (avg. overvaluation during a run above fundamentals), %	7.9	6.7	9.0	8.5
Lower semi-deviation (avg. undervaluation during a run below fundamentals), %	8.9	1.6	1.8	8.6
Average volatility (per trading round), %	2.9	2.2	3.6	2.8
Behavioural and budget constraints				
Average proportion of agents forming "unreasonable" dividend forecast (per forecasting round), %	0.0	0.0	5.0	0.0
Average number proportion of agents that have "unreasonable" reservation price (per trading round), %	0.1	0.4	3.3	0.1
Number of active agents, %	89.7	29.2	22.2	90.5
Adaptive adjustment				
Average dividend adjustment factor	1.015	1.016	0.954	0.998
Average price adjustment factor	0.986	1.004	0.973	1.002


6000
5000
4000
3000
2000
1000
0
5000
10000
15000
20000
Market price
Perceived fundamental price
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
10
20
30
40
50
60
70
80
90
100
Individual wealth level (at the end of a typical simulation run)
400
300
200
100
0
5000
10000
15000
20000
Actual dividend realisation (ex post)
5-year dividend forecast (superimposed)
100
80
60
40
20
0
5000
10000
15000
20000
Percentage of agents breaching imposed constraints on reservation price
100
80
60
40
20
0
0
5000
10000
15000
20000
Percentage of active agents
0.08
0.06
0.04
0.02
0
5000
10000
15000
20000
Market price volatility (per trading period)

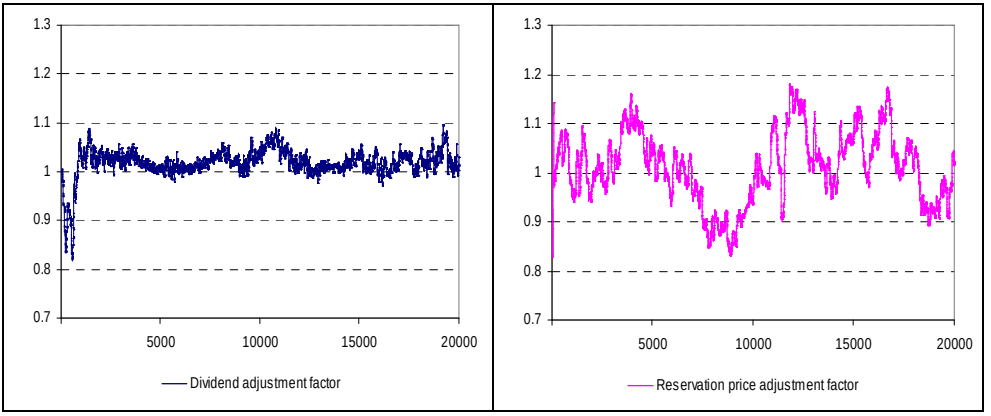
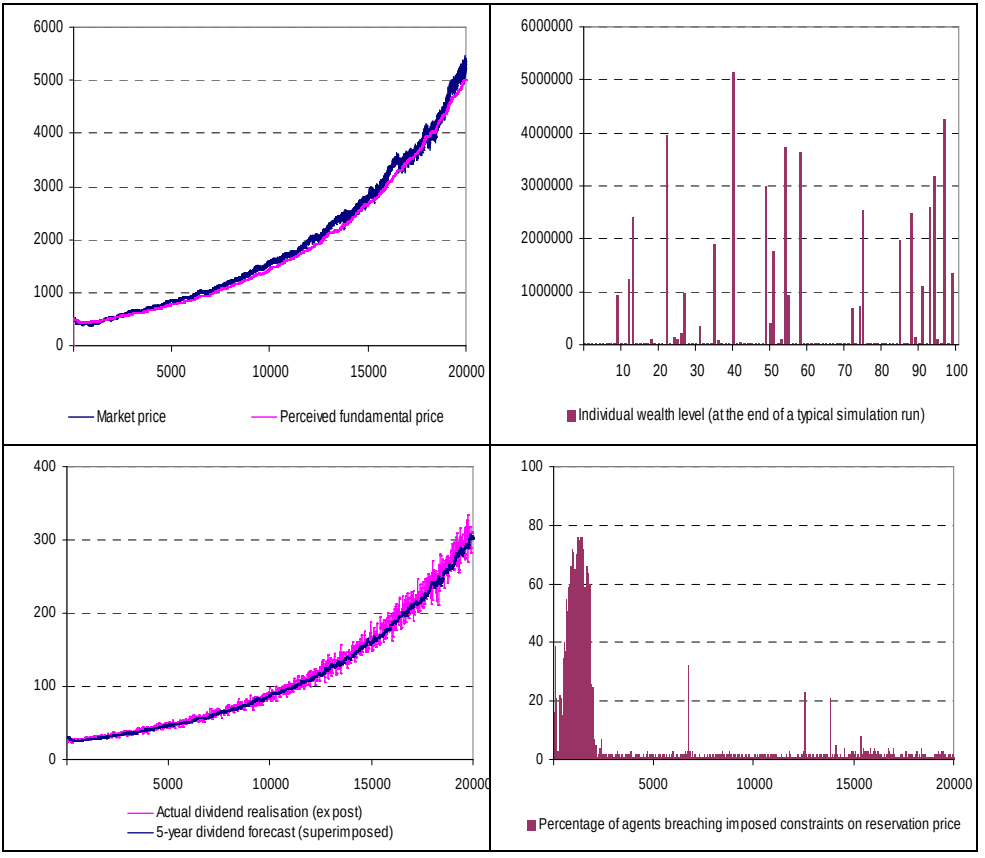


Fig. A1. Selected graphs of Experiment 1



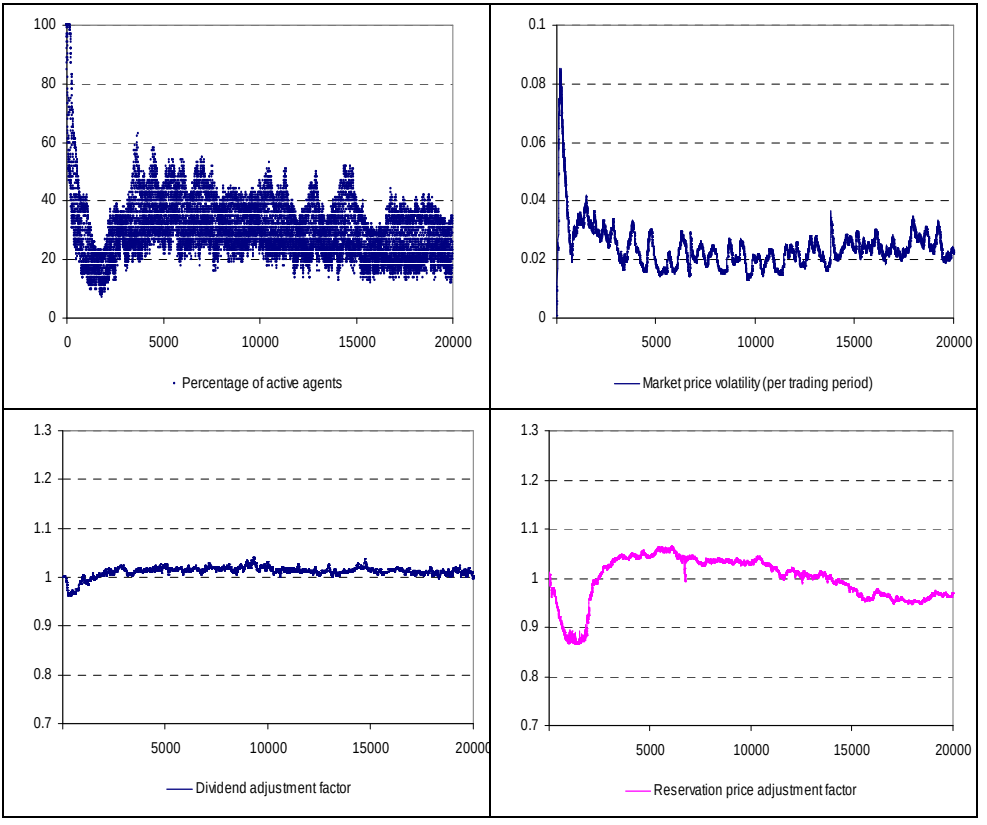
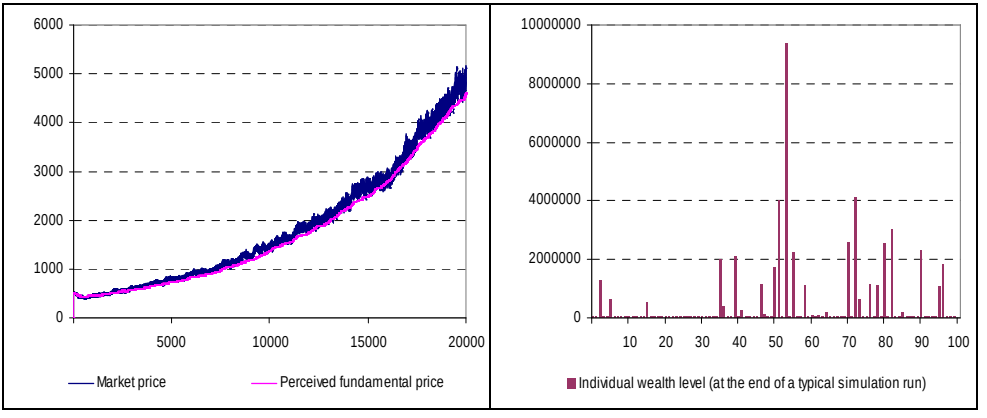


Fig. A2. Selected graphs of Experiment 2



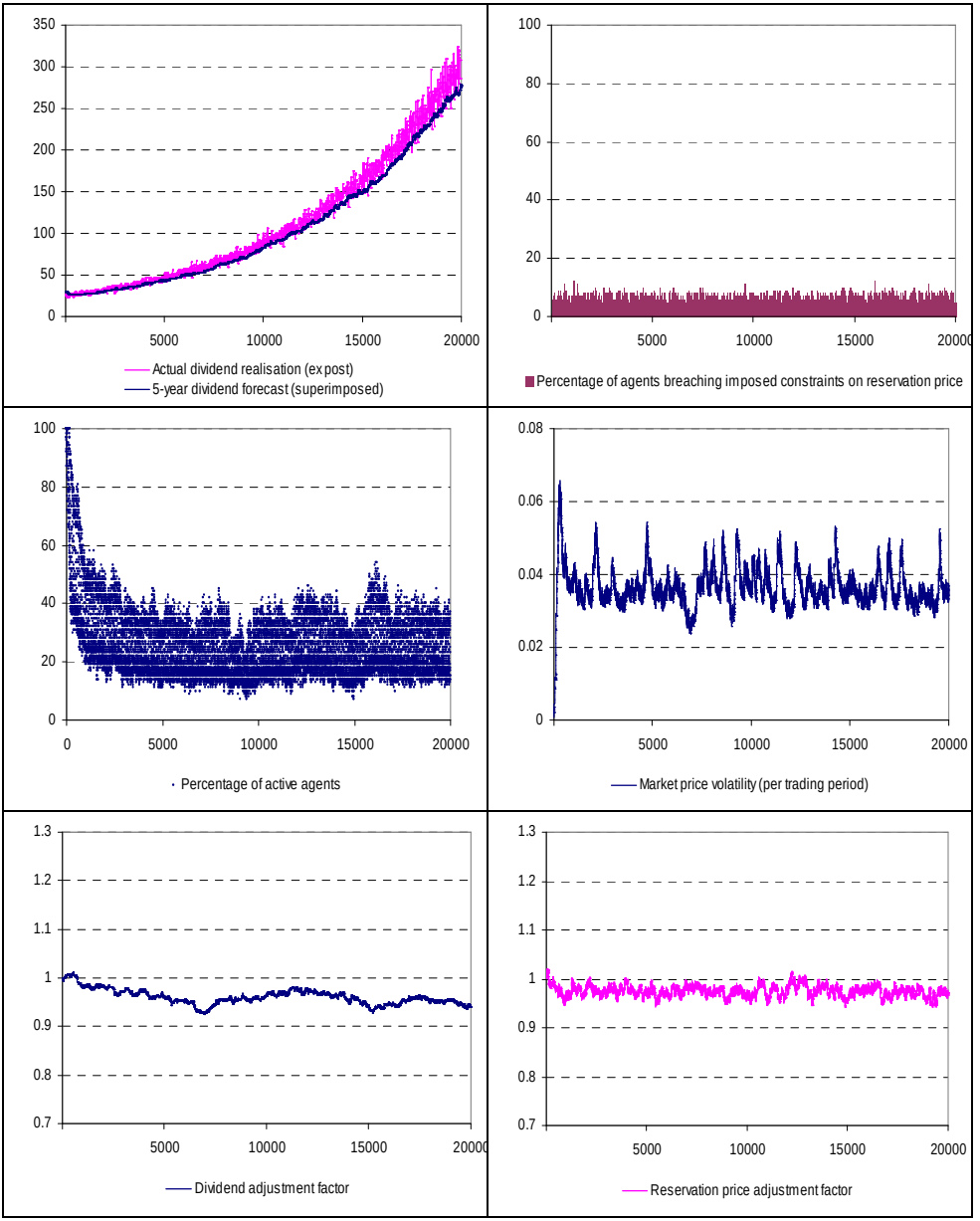
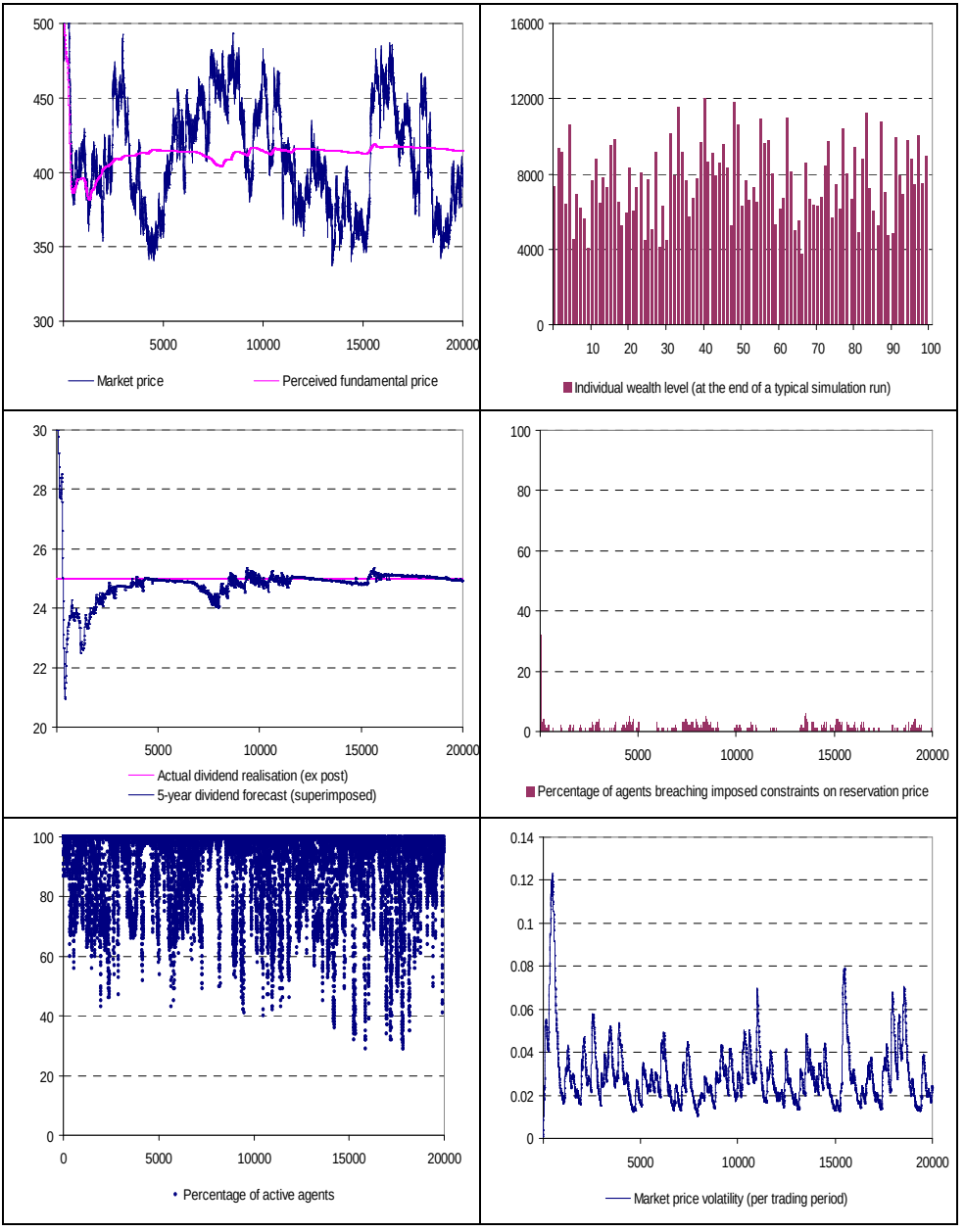


Fig. A3. Selected graphs of Experiment 3


500
450
400
350
300
5000
10000
15000
20000
Market price
Perceived fundamental price
16000
12000
8000
4000
0
10
20
30
40
50
60
70
80
90
100
Individual wealth level (at the end of a typical simulation run)
30
28
26
24
22
20
5000
10000
15000
20000
Actual dividend realisation (ex post)
5-year dividend forecast (superimposed)
100
80
60
40
20
0
5000
10000
15000
20000
Percentage of agents breaching imposed constraints on reservation price
100
80
60
40
20
0
0
5000
10000
15000
20000
Percentage of active agents
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
5000
10000
15000
20000
Market price volatility (per trading period)

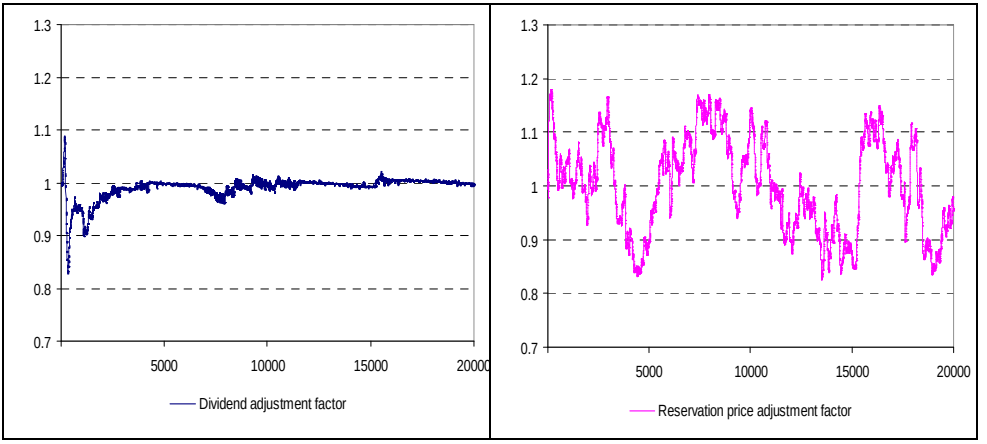


Fig. A4. Selected graphs of Experiment 4

Appendix B. Parameter setting and experimental results of parsimonious version of the model

Table B1. Model parameters

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 (alpha)	0.1
Exploration rate (epsilon)	0.1
Subjective discount parameter of reinforcement learning	0.995
Action (incremental adjustment in equation (4))	-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

Table B2. Experimental setting in the case of stationary exogenous processes

Simulation size	
Length of simulation run (trading rounds *t*)	15000
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

Table B3. Experimental setting in the case of calibration to actual data

Simulation size			
Length of a simulation run (trading rounds *t*)	5290		
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

Table B4. Properties of simulated and actual returns

	Actual	**Experiment**		
	SP 500	**1**	**2**	**3**
Penalty rate		-2	-3	-3
Loss aversion (%)		0	0	10
Initial sample				
Number of observations	4291	4291	4291	4291
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	4291	4267	4268	4263
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	19459.71	16080.79	18253.55	14013.62
Probability	0.000000	0.000000	0.000000	0.000000
1-lag autocorrelation of returns	-0.076	0.044	0.076	0.070
1-lag autocorrelation of sqr. 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

Appendix C. Program code for the initial version of the model

Main program file

```
% Artificial stock market (ASM), populated by heterogeneous agents. Their behaviour is based on Q-learning
% with gradient-descent function approximation. Program implementation in MATLAB version R2007a.

clear all

% Model's key parameters
stock_p     = 500;                                          % Initial stock price
iterations  = 20000;                                        % Number of iterations
agents      = 100;                                          % Number of agents
num_shares  = 10000;                                        % Number of shares
grid_p      = 50;                                           % Number of possible price choices
bond_p      = 0.995;                                        % Bond price
alpha       = 0.1;                                          % Learning rate parameter
epsilon     = 0.1;                                          % Exploration parameter
trade_cost  = 0.001;                                        % Trade cost (as a fraction of transacted price)
for_horizon = 5;

% Technical initialisation
sigma_p  = 10;                                              % Initial std. dev. of stock price
dividend = 25;                                              % Initial dividend
periods  = 12;                                              % No. of periods between dividend payouts
next_div = periods - 1;                                     % Time periods until next dividend payout
h        = zeros(agents,1) + num_shares / agents;           % Stock holdings before trading round
h1       = h;                                               % Stock holdings after trading round
m        = zeros(agents,1) + 700;                           % Cash holdings before trading round
m1       = m;                                               % Cash holdings after trading round
wealth   = h1 * stock_p + m1;                               % Individual wealth at market prices
exp_sigma_p   = sigma_p;                                    % Initial expectation of std. dev. of stock price
stock_p_prev  = stock_p;                                    % Last period's market stock price
perform       = zeros(agents,1);                            % Performance measure
```

```
std_err       = 0;                                    % Smoothed standard deviation of dividend process
volatility    = 0;                                    % Stock price volatility {state2.m}
bound1        = 0.3;                                  % Imposed constraint on dividend forecasts
bound2        = 0.2;                                  % Imposed constraint on stock price estimates
adjust1       = ones(agents,1);                       % RL adjustment vector, dividend forecast case
adjust2       = ones(agents,1);                       % RL adjustment vector, stock price case
exp_fund      = 500;                                  % Unadjusted fundamental stock value
exp_div = zeros(agents,1) + 30 + rand(agents,1) ...   % Initial dividend forecast (5-step)
... * 5 - rand(agents,1) *5;
exp_p_LR   = exp_div / ((1/bond_p) ^ periods - 1);    % Initial adjusted fundamental stock value
avg_p       = stock_p + rand * 5;                     % Initiation value of EWMA of market stock price
explore       = zeros(agents,1);                      % Indicator variable (explore vs. greedy action)
buy_prob      = ones(grid_p,1);                       % Prob. of buy order execution (at given price)
sell_prob     = ones(grid_p,1);                       % Prob. of sell order execution (at given price)
prob          = ones(grid_p,2);                       % Matrix of successful trade probs (buy,sell)
split         = zeros(agents,1);                      % Indicator variable (split of best-performers)

% Plotted variables and analytical indicators
plot_p_LR = zeros(iterations,1);
plot_p = zeros(iterations,1);
plot_div = ones(iterations,2) * 30;
plot_fund = zeros(iterations,1);
plot_rewards2  = zeros(iterations,1);
plot_p_LR_active = zeros(iterations,1);
plot_fund_active = zeros(iterations,1);
plot_active = zeros(iterations,1);
plot_turnover = zeros(iterations,1);
plot_volatility = zeros(iterations,1);
plot_adjust1 = zeros(iterations,1);
plot_adjust2 = zeros(iterations,1);
plot_money = zeros(iterations,1);

% State features in the linear approximation of Q-value
features1      = zeros(agents,4);                     % Matrix of state features, dividend case
features2      = zeros(agents,4);                     % Matrix of state features, stock price case
f_history = zeros(agents,size(features1,2)*(for_horizon+1));  % History of state features, dividend case
a_history = zeros(agents,for_horizon+1) + 2;          % History of actions, dividend case
div_forecast = ones(agents,for_horizon+1) * 30;       % Matrix of agents' dividend forecasts
div_history = ones(agents,for_horizon+1) * 30;        % Matrix of EWMAs of dividend process
```

```
% Experience matrices of Q-learning algorithm
theta1         = zeros(size(features1,2) * agents, 3);
theta2         = zeros(size(features2,2) * agents, 3);

% Agents randomly take initial actions
grid_a = [ceil(rand(agents,1) * 3) ceil(rand(agents,1) * 3)];  % Agents' action matrix (dividend & price
                                                               % adjustment variables)

% Main cycle
for k = 1 : iterations
    next_div = next_div - 1;

    % Update and store values of financial accounts (before trading round)
    lqd_need = stock_p * 5;
    h_prev = h;
    h = h1;
    m_prev = m;
    m = m1 / bond_p;
    m(m > lqd_need) = lqd_need;
    m_average = sum(m) / length(m);

    % Form expectations, submit orders, trade and record trading statistics
    [alpha1,alpha2,exp_div,div_forecast,div_history,adjust1,exp_fund,adjust2,exp_p_LR] = ...
    ... expectations(grid_a,dividend,exp_div,div_forecast,div_history,for_horizon,next_div, ...
    ... periods,bond_p,adjust1,exp_fund,adjust2);
    [orders,exp_sigma_p,prob,p_quotes] = make_orders(grid_p,stock_p,exp_p_LR,exp_sigma_p, ...
    ... agents,h,m,trade_cost, div_forecast,next_div,buy_prob,sell_prob,prob,explore);
    [orders_e,stock_p,stock_p_prev,turnover,transactions,sigma_p,lqd_need,buy_prob,sell_prob] = ...
    ... trade(orders,stock_p,stock_p_prev,sigma_p,lqd_need,grid_p,p_quotes,prob);

    % Stochastic exogenous realisation of dividends on regular time intervals
    if next_div == 0
        dividend = 25 * 1.000125 ^ k + 0.05 * plot_div(k-1,2) * randn;
        plot_div(k,:) = [mean(div_forecast(:,end)) dividend];
    else
        dividend = 0;
        if k>1, plot_div(k,:) = plot_div(k-1,:); end
    end
```

```
% Update investors' accounts (after trading round)
bs = orders_e(:,3);                                          % Indicator variable (buy or sell)
h1 = h + orders_e(:,2) .* bs;
m1_prev = m1;
m1 = m - orders_e(:,2) .* orders_e(:,1) .* bs ...            % Dividend is paid after trading round,
... - orders_e(:,2) .* orders_e(:,1) * trade_cost + h1 * dividend;            % hence h1*dividend
wealth_prev = wealth;
wealth = h1 * stock_p + m1;

% Observe resulting states and rewards, dividend forecasting case
if next_div == 0
    [features1,f_history,f_prev1,std_err] = ...
    ... state1(features1,f_history,div_history,for_horizon,dividend,std_err,adjust1);
    rewards1 = - (dividend - div_forecast (:,end) ) .^ 2 / 100;    % Squared forecast error
    % rewards1 = zeros(agents,1);

    avg_fcst_error(k) = mean((dividend - div_forecast(:,end)) ./ div_forecast(:,end));
    avg_abs_error(k) = mean(abs((dividend - div_forecast(:,end)) ./ div_forecast(:,end)));

    % Impose 'reasonable' limits on dividend forecasts
    limit1 = find (div_forecast(:,1) < (1 - bound1) * dividend);
    limit2 = find (div_forecast(:,1) > (1 + bound1) * dividend);
    if ~isempty(limit1) && k > for_horizon * periods
        rewards1(limit1) = rewards1(limit1) * 2;
        adjust1(limit1) = 1.01 * adjust1(limit1);
    end
    if ~isempty(limit2) && k > for_horizon * periods
        rewards1(limit2) = rewards1(limit2) * 2;
        adjust1(limit2) = 0.99 * adjust1(limit2);
    end
end

% Observe resulting states and rewards, stock value estimation case
[f_prev2,features2,avg_p,volatility] = state2 (features2,agents,stock_p,avg_p,volatility,adjust2);
rewards2 = (log((h1 * stock_p + m1 / bond_p) - log(h * stock_p_prev + m)));    % Log-returns
rewards2 = rewards2 + rewards2 .* split * 0.5;
%rewards2 = zeros(agents,1);
```

```
% Impose 'reasonable' limits on stock value estimates
limit1 = find(exp_p_LR < (1 - bound2) * exp_fund);
limit2 = find(exp_p_LR > (1 + bound2) * exp_fund);
    if ~isempty(limit1)
        rewards2(limit1) = rewards2(limit1) - abs(rewards2(limit1)) * 2;
        adjust2(limit1) = adjust2(limit1) * 1.02;
        out_of_bounds2(k) = length(limit1);
    end
    if ~isempty(limit2)
        rewards2(limit2) = rewards2(limit2) - abs(rewards2(limit2)) * 2;
        adjust2(limit2) = adjust2(limit2) * 0.98;
        out_of_bounds2(k) = out_of_bounds2(k) + length(limit2);
    end

explore = zeros(agents,1);

% Implement learning procedure for each individual investor and take action
for i=1:agents
    % Individual i's action reindexing
    action_index1 = a_history(i,end);                          % Dividend forecasting case
    action_index2 = grid_a(i,2);                               % Price forecasting case
    agent_index1 = (((i-1) * size(features1,2)+ 1):((i-1) ... % Determine 'experience' matrix ...
    ... * size(features1,2) + size(features1,2)))';            % row indices associated with i-th investor
    agent_index2 = (((i-1) * size(features2,2)+ 1):((i-1) * size(features2,2) + size(features2,2)))';

    % Learning procedure
    if next_div == 0
        [theta1,max_ind1] = learn (theta1,features1,f_prev1,agent_index1,action_index1,rewards1,alpha,i);
        % Learn to forecast dividends, 'last period' in this case is t - forecast horizon
    end
    [theta2,max_ind2] = learn (theta2,features2,f_prev2,agent_index2,action_index2,rewards2,alpha,i);

    % Decide on next action and take it
    if rand > epsilon                              % If it is the exploitation phase...
        if next_div == 0                           % ... and if it's the dividend payout period...
            grid_a(i,1) = max_ind1;                % ... then take greedy action (dividend forecasting)
        end
        grid_a(i,2) = max_ind2;                    % ... take greedy action (stock value estimation)
    else                                           % ... otherwise take random exploratory actions
```

```
        if next_div == 0
            grid_a(i,1) = ceil(rand * 3);
        end
        grid_a(i,2) = ceil(rand * 3);
        explore(i) = 1;
    end

end

% Update action history matrix (with current action stored in the left-most column)
if next_div == 0
    for j = 1:for_horizon
        z = for_horizon + 2 - j;
        a_history(:,z) = a_history(:,z-1);
    end
    a_history(:,1) = grid_a(:,1);
end

% Strategy imitation and evolutionary selection of agents
perform = perform + rewards2;
[theta1,theta2,adjust1,adjust2,perform,m1,h1,split] = ...
... evolution(perform,agents,features1,features2,theta1,theta2,adjust1,adjust2,m1,h1);
%[theta2, perform, adjust2] = imitation(theta2, perform, adjust2, features2, agents);      % Optional
%[theta1,theta2,adjust1,adjust2,perform,m1,h1] = ...                                       % Optional
%... noise_traders(perform,agents,features1,features2,theta1,theta2,adjust1,adjust2,m1,h1);

if next_div == 0, next_div = periods; end

% Update variables used for analytical and illustrative purposes
plot_p(k) = stock_p;
plot_p_LR(k) = sum(exp_p_LR) / agents;
plot_fund(k) = sum(exp_fund) / agents;
owns_m = m > stock_p;
owns_stock = h > 1;
active = owns_m .* owns_stock;
plot_p_LR_active(k) = sum(exp_p_LR .* active) / sum(active);
plot_fund_active(k) = sum(exp_fund .* active) / sum(active);
plot_active(k) = sum(active);
plot_rewards2(k) = sum(rewards2) / agents;
```

```
    plot_turnover(k) = turnover;
    plot_money(k) = sum(m1);
    plot_volatility(k) = volatility / stock_p;
    plot_adjust1(k) = sum(adjust1) / length(adjust1);
    plot_adjust2(k) = sum(adjust2) / length(adjust2);

    % Arbitrary program execution indicators
    [k stock_p plot_p_LR(k) plot_fund(k) plot_div(k,2) n]

end

% Implement procedure of storing data on file (not reported here)
<…>
```

Function describing formation of individual expectations

```
% This function computes agents' individual expectations regarding dividend prospects and determines
% individual estimates of stock price fundamentals. It relies on results of reinforcement learning.

function [alpha1,alpha2,exp_div,div_forecast,div_history,adjust1,exp_fund,adjust2,exp_p_LR] = ...
... expectations(grid_a,dividend,exp_div,div_forecast,div_history,for_horizon,next_div,periods, ...
... bond_p,adjust1,exp_fund,adjust2)

gamma   = 0.1;                                          % Smoothing factor in EWMA

% Changes in dividend and price adjustment variables (determined by Q-learning
alpha1  = (grid_a(:,1) - 2) / 50;
alpha2  = (grid_a(:,2) - 2) / 50;

% Dividend forecast procedure, which takes place following dividend payout
if next_div == periods - 1
    exp_div = (1 - gamma) * exp_div + gamma * dividend;    % Exponential moving average of dividend payouts

    % Generating dividend forecast and updating matrices (dividend history and dividend forecasts)
    for i = 1 : for_horizon
        z = for_horizon + 2 - i;
        div_forecast(:,z) = div_forecast(:,z-1);
        div_history(:,z) = div_history(:,z-1);
```

```
    end
    adjust1 = adjust1 .* (1 + alpha1);
    div_forecast(:,1) = exp_div .* adjust1;
    div_history(:,1) = exp_div;

    % Calculate individual estimates of 'raw' fundamental stock value from expected discounted income flows
    i_rate = (1/bond_p) ^ periods - 1;             % Annual interest rate calculated from the bond price
    fundament = dividend / 2;
    for i=1:for_horizon
        fundament = fundament + div_forecast(:,i) ./ ((1 + i_rate) ^ (for_horizon - i + 1));
    end
    fundament = fundament + (div_forecast(:,1) + div_forecast(:,1) / i_rate) ... % By assumption, dividend
    ... / ((1 + i_rate) ^ (for_horizon + 1));                % becomes constant beyond forecasting horizon
    exp_fund = (1- gamma) * exp_fund + gamma * fundament;
end

% Calculate individual estimates of 'adjusted' fundamental stock value
adjust2 = adjust2 .* (1 + alpha2);
exp_p_LR = exp_fund .* adjust2;
```

Fuction describing formation of buy/sell orders

```
% This function transforms the indices of taken actions into actual buy/sell orders.

function [orders,exp_sigma_p,prob,p_quotes] = ...
... make_orders(grid_p,stock_p,exp_p_LR,exp_sigma_p,agents,h,m,trade_cost,div_forecast, ...
... next_div,buy_prob,sell_prob,prob,explore)

% Technical assumptions and initialisation
gamma = 0.1;                        % Smoothing factor
orders = zeros(agents,4);           % Matrix containing price, quantity, buy/sell indication, agent ID
sigma_p = 0.05 * stock_p;           % Assumed initial price volatility
exp_sigma_p     = gamma * sigma_p + (1 - gamma) * exp_sigma_p;  % Smoothed price volatility
p_quotes = zeros(grid_p,1);                                     % Vector containing feasible price quote grid

% Partitioning of standard normal distribution for the case of price quote grid of 50
std_norm = [ 0 0.0502 0.1004 0.1510 0.2019 0.2533 0.3055 0.3585 0.4125 0.4677 0.5244 0.5828 0.6433 ...
    0.7063 0.7722 0.8416 0.9154 0.9945 1.0803 1.1750 1.2815 1.4051 1.5548 1.7507 2.0538 3.0905 ];
```

```
% Calculating demand and supply functions from trading statistics
prob(:,1) = gamma * buy_prob + (1 - gamma) * prob(:,1);   % Matrix containing smoothed probabilities of
prob(:,2) = gamma * sell_prob + (1 - gamma) * prob(:,2);  % buying/selling a share at a all possible prices

prob(end,1) = 1;                                          % Agents assume that they can always buy/sell
prob(1,2) = 1;                                            % stocks at marginal prices
y = (1:grid_p)';
polynom1 = polyfit(y,prob(:,1),1);                        % Fit demand and supply lines
polynom2 = polyfit(y,prob(:,2),1);
prob(:,1) = polyval(polynom1,y);
prob(:,2) = polyval(polynom2,y);
prob(prob>1) = 1;                                         % Constrain probabilities to interval [0,1]
prob(prob<0) = 0;

% Constructing feasible price quote grid, which is more dense around the prevailing market price
for i=1:(grid_p/2)
    p_quotes(i) = (-std_norm(grid_p / 2 - i + 1) * exp_sigma_p ...
    ... + stock_p - std_norm(grid_p / 2 - i + 2) * exp_sigma_p + stock_p) / 2;
    p_quotes(grid_p + 1 - i) = 2 * stock_p - p_quotes(i);
end

% Evaluating alternatives and choosing trade orders for all agents
for k=1:agents

    % Determine price range, in which a given agent is willing to trade
    exp_q = zeros(grid_p,1);     % Vector containing exp. prob to sell/buy at all prices for a given agent
    if stock_p < exp_p_LR(k)     % If mkt. price is lower than perceived value then agent wants to buy stock
        min_i = min(find(p_quotes > 0));
        max_i = max(find(p_quotes < exp_p_LR(k)));
        bs = 1;                                       % Variable indicating that agent wants to buy/sell stock
    else
        min_i = min(find(p_quotes > exp_p_LR(k)));
        max_i = grid_p;
        bs = - 1;
    end

    % Put demand/supply restrictions (sufficient cash, stock holdings, max trade of 1 unit)
    max_q = zeros(grid_p,1);
```

```
for i = min_i : max_i
    if bs == 1 && m(k) > (1 + trade_cost) * p_quotes(i);
        max_q(i) = 1 ;
    elseif bs == -1 && h(k) > 1
        max_q(i) = 1;
    else
        max_q(i) = 0;
    end
end

% Check whether dividend is being paid out in the current period
if next_div == 0
    e_div = div_forecast(k,end);
else
    e_div = 0;
end

% Choose transaction price which gives highest expected end-of-period wealth
if ~explore(k)                                  % If agent is taking a greedy action...

    % Determine expected traded quantity
    if bs == 1
        exp_q = prob(:,1) .* max_q;
    else
        exp_q = prob(:,2) .* max_q;
    end

    % Determine expected wealth after trading round
    exp_h1 = h(k) + exp_q * bs;
    exp_m1 = m(k) - exp_q .* p_quotes * bs - exp_q .* p_quotes * trade_cost + exp_h1 * e_div;
    exp_wealth = exp_h1 * exp_p_LR(k) + exp_m1;

    % Choose wealth-maximising price
    max_wealth_ind = find(exp_wealth == max(exp_wealth));
    x = randperm(length(max_wealth_ind));
    optimal_i = max_wealth_ind(x(1));

else                                             % ... otherwise, if agent is taking exploratory action...
    optimal_i = min_i + ceil(rand * (max_i - min_i)); % Choose price randomly from available price quotes
```

```
    end

    % Fill a row associated with the given agent's order  in the orders matrix
    orders(k,1) = p_quotes(optimal_i);
    orders(k,2) = max_q(optimal_i);
    orders(k,3) = bs;
    orders(k,4) = k;
end
```

Function describing the trading process

```
% This function implements the trading process, assuming a simplified version of the double auction (DA)

function [orders_e,stock_p,stock_p_prev,turnover,transactions,sigma_p,lqd_need,buy_prob,sell_prob] = ...
... trade(orders,stock_p,stock_p_prev,sigma_p,lqd_need,grid_p,p_quotes,prob)

% Initialisation and variable desription
n = length(orders(:,1));                                        % Number of orders to be processed
buy_prob = zeros(size(p_quotes));
sell_prob = zeros(size(p_quotes));

% Queue orders in a random manner
orders_q = sortrows([orders rand(n,1)],5);
orders_q(:,5) = [];
orders_e = [zeros(n,2) orders_q(:,3:4)];

% Order matching and execution procedure
for i = 2:n
    opposite = orders_q(i,3) * (-1);                            % Determine opposite order (buy or sell)
    complete = false;                                           % Logical variable, (executed order)
    pending = false;                                            % Logical variable (pending order)
    if orders_q(i,2) == 0, complete = true; end

    % Process a given order until it gets status “executed” of “pending”
    while ~(complete || pending)

        %Search for potential matches
        if opposite == -1                                       % Search for sell orders
```

```
            match_con = (orders_q(1:(i-1),3) == opposite) .* ...
            ... (orders_q(1:(i-1),1) <= orders_q(i,1)) .* (orders_q(1:(i-1),2) > 0);
        else                                                    % Search for buy orders
            match_con = (orders_q(1:(i-1),3) == opposite) .* ...
            ... (orders_q(1:(i-1),1) >= orders_q(i,1)) .* (orders_q(1:(i-1),2) > 0);
        end

        % Find best matches and execute orders
        if max(match_con) > 0                                   % Check if there are suitable opposite orders

            % Find best matches
            orders_p = orders_q(1:(i-1),:) .* repmat(match_con,1,4); % List of potential orders
            if opposite == -1
                ask_p = orders_p(:,1);                          % List of potential ask prices
                p_best = min(ask_p(ask_p > 0));                 % Lowest positive ask price
                I = find(ask_p == p_best, 1);                   % Determine that order's index
            else
                [p_best I] = max(orders_p(:,1));                % Index of best buy order
            end

            % Execute trade and update pending order list
            price = (p_best + orders_q(i,1)) / 2;               % Trade at average of bid and ask prices
            volume = min(orders_q(I,2),orders_q(i,2));          % volume = 1 in this model version
            orders_q(I,2) = orders_q(I,2) - volume;             % Update pending order list
            orders_q(i,2) = orders_q(i,2) - volume;

            % Update executed orders' list
            orders_e(I,1) = (orders_e(I,1)* orders_e(I,2)+ price * volume) / (orders_e(I,2) + volume);
            orders_e(I,2) = orders_e(I,2) + volume;
            orders_e(i,1) = (orders_e(i,1)* orders_e(i,2)+ price * volume) / (orders_e(i,2) + volume);
            orders_e(i,2) = orders_e(i,2) + volume;
            if orders_q(i,2) == 0, complete = true; end
        else                                                    % If no potential matches, pending order
            pending = true;
        end
    end
end

% Calculate trading round statistics (avg stock price, turnover, number of transactions)
```

```
turnover = 0;
transactions = 0;
orders_e = sortrows(orders_e,4);
if sum(orders_e(:,2)) > 0.001                                    % If non-zero turnovers are attained...
    stock_p_prev = stock_p;                                      % Store last period's stock price
    stock_p = sum(orders_e(:,1) .* orders_e(:,2) / sum(orders_e(:,2)));  % Trading round's (average) price
    turnover = sum(orders_e(:,1) .* orders_e(:,2)) / 2;          % Turnover (in monetary terms)
    transactions = sum(orders_e(:,2)>0.0001) / 2;                % Number of transactions
end

% Compute successful trade probabilities
for i=1:grid_p                                                   % For each price quote...
    x = find(orders(:,1) == p_quotes(i));                        % ID of agents that used this quote
    buy = x(orders(x,3) == 1);                                   % ID of agents that wanted to buy at this price
    sell = x(orders(x,3) == -1);                                 % ID of agents that wanted to sell at this price
    if ~isempty(buy)
        buy_prob(i) = sum(orders_e(buy,2)) / length(buy); % Prob. of successful buy order at this price
    else
        buy_prob(i) = prob(i,1);                                 % Otherwise, previous estimate retained
    end

    if ~isempty(sell)
        sell_prob(i) = sum(orders_e(sell,2)) / length(sell);
    else
        sell_prob(i) = prob(i,2);
    end
end
```

Function describing state variables (dividend forecasting case)

```
% This function computes matrix containing features of the state1 (dividend forecasting case)

function [features1,f_history,f_prev1,std_err] = ...
... state1 (features1,f_history,div_history,for_horizon,dividend,std_err,adjust1)


gamma = 0.1;                                                     % Smoothing parameter in EWMA
sq_dev = (dividend - div_history(:,1)) .^ 2;                     % Squared deviation from exponential moving average
```

```
std_err = sqrt ( gamma * sq_dev + (1 - gamma) * std_err .^ 2 );   % Exponentially weighted standard deviation

features1(:,1) = (adjust1 - 1);                                   % Current adjustment factor
features1(:,2) = (dividend - div_history(:,1)) ./ std_err / 10; % Deviation from avg., relative to std. dev.
features1(:,3) = features1(:,2) .^  2;                            % Second feature squared
features1(:,4) = dividend ./ div_history(:,1) / 10;               % Dividend, as compared to the EWMA

% Store previous realisations of state features in a matrix
x = size(features1,2);
for i=1:for_horizon
    z = for_horizon + 2 - i;
    f_history(:,(z-1)*x+1:z*x) = f_history(:,(z-2)*x+1:(z-1)*x);
end
f_history(:,1:x) = features1;
f_prev1 = f_history(:,end-x+1:end);
```

Function describing state variables (value estimation case)

```
% This function computes matrix containing features of the state (stock value estimation case)

function [f_prev2,features2,avg_p,volatility] = state2(features2,agents,stock_p,avg_p,volatility,adjust2)

f_prev2 = features2;                                    % Store previous values of feature matrix
features2 = zeros(agents,4);
gamma = 0.01;

avg_p = (1 - gamma) * avg_p + gamma * stock_p;          % EWMA of stock price
sq_dev = (stock_p - avg_p) .^ 2;                        % Squared deviation from EWMA
volatility = sqrt (gamma * sq_dev + (1 - gamma) * volatility .^ 2); % Exponentially weighted std. deviation

features2(:,1) = (adjust2 - 1);                         % Current adjustment factor
features2(:,2) = (stock_p - avg_p) ./ volatility / 10; % Deviation from avg., relative to std. deviation
features2(:,3) = features2(:,2) .^ 2;                   % Second feature squared
features2(:,4) = stock_p ./ avg_p / 10;                 % Current stock price, as compared to EWMA
```

Function describing the learning procedure

```
% This function implements reinforcement learning procedure

function [theta,max_ind] = learn (theta,features,f_prev,agent_index,action_index,rewards,alpha,i)

gamma = 0.995;                                              % RL discount rate

Q_s_a = f_prev(i,:) * theta(agent_index,action_index);      % Estimate of Q-function
delta = rewards(i) - Q_s_a;
Q_s1_all = features(i,:) * theta(agent_index,:);            % All estimates of Q-functions of resultant state
max_Q_ind = find(Q_s1_all == max(Q_s1_all));                % Action indices of the largest Q-functions
x = randperm(length(max_Q_ind));                            % Perform random permutation...
max_ind = max_Q_ind(x(1));                                  % and choose any of optimum Q-functions
max_Q = Q_s1_all(max_ind);                                  % Find maximum Q-function of resultant state
delta = delta + gamma * max_Q;                              % Add discounted Q-function to delta
theta(agent_index,action_index) = ...                       % Update experience matrix theta
...theta(agent_index,action_index) + alpha * delta * f_prev(i,:)';
```

Function governing evolutionary selection of agents

```
% This function models exit of worst-performing agents and increasing prevalence of most successful agents

function [theta1,theta2,adjust1,adjust2,perform,m1,h1,split] = ...
... evolution(perform,agents,features1,features2,theta1,theta2,adjust1,adjust2,m1,h1)

    split = zeros(agents,1);            % Vector indicating which successful agent replaced
                                        % least successful ones (and split their holdings)

% Rate investors according to their performance (cumulative rewards2)
rating = [ perform (1:agents)'];
lead = sortrows(rating,-1);             % Winners (performance, ID)
lose = sortrows(rating,1);              % Losers (performance, ID)

% Investors exit the market if wealth falls below arbitrary threshold with respect to average wealth level
avg_perform = sum(perform) / agents;
bankrupt = 3;                       % Model-imposed number of forced bankruptcies
```

```
% Most successful agents take worst-performing agents' place
if ~isempty(sum(perform < 0.7 * avg_perform))
   for i=1:bankrupt
      % Determine indices of most and least successful agents
      agent_index_lead1 = (((lead(i,2)-1) * size(features1,2)+ 1):((lead(i,2)-1) ...
      ... * size(features1,2) + size(features1,2)))';
      agent_index_lose1 = (((lose(i,2)-1) * size(features1,2)+ 1):((lose(i,2)-1) ...
      ... * size(features1,2) + size(features1,2)))';
      agent_index_lead2 = (((lead(i,2)-1) * size(features2,2)+ 1):((lead(i,2)-1) ...
      ... * size(features2,2) + size(features2,2)))';
      agent_index_lose2 = (((lose(i,2)-1) * size(features2,2)+ 1):((lose(i,2)-1) ...
      ... * size(features2,2) + size(features2,2)))';

      % Leaders take over least successful agents' financial holdings and split into two
      h1(lead(i,2)) = (h1(lose(i,2)) + h1(lead(i,2))) / 2;
      m1(lead(i,2)) = (m1(lose(i,2)) + m1(lead(i,2))) / 2;
      h1(lose(i,2)) = (h1(lose(i,2)) + h1(lead(i,2))) / 2;
      m1(lose(i,2)) = (m1(lose(i,2)) + m1(lead(i,2))) / 2;

      % Sharing experience of successful wealth management
      theta1(agent_index_lose1,:) = theta1(agent_index_lead1,:);
      theta2(agent_index_lose2,:) = theta2(agent_index_lead2,:);
      adjust1(lose(i,2)) = adjust1(lead(i,2));
      adjust2(lose(i,2)) = adjust2(lead(i,2));
      perform(lose(i,2)) = perform(lead(i,2));
      split(lead(i,2)) = 1;
   end
end
```

Function determining imitation of strategies

```
% This function models sharing stock trading experience, whereby randomly matched investors imitate
% investment strategy of more successful counterparty

function [theta2,perform,adjust2] = imitation(theta2,perform,adjust2,features2,agents)

% Random matching of agents
rand_match = [(1:agents)' rand(agents,1)];
```

```
rand_pairs = [(1:agents)' sortrows(rand_match,2)];
rand_pairs(:,3) = [];

for i=1:agents
    perform_pair = [ perform(rand_pairs(i,1)) perform(rand_pairs(i,2)) ];     % Performance measures
    [perform_better I_better] = max(perform_pair);                            % Identify more successful one
    [perform_worse I_worse] = min(perform_pair);                              % Identify less successful one

    % Locate appropriate portions of paired agents' experience matrices
    agent_index_b = (((rand_pairs(i,I_better)-1) * ...
    ... size(features2,2)+ 1):((rand_pairs(i,I_better)-1) * size(features2,2) + size(features2,2)))';
    agent_index_w = (((rand_pairs(i,I_worse)-1) * ...
    ... size(features2,2)+ 1):((rand_pairs(i,I_worse)-1) * size(features2,2) + size(features2,2)))';

    % Less successful agent imitates with prespecified probability
    if (perform_better - perform_worse) / abs(perform_better) > rand * 0.2
        theta2(agent_index_w,:) = theta2(agent_index_b,:);
        perform(I_worse) = perform(I_better);
        adjust2(I_worse) = adjust2(I_better);
    end
end
```

Function describing noise trading behaviour

```
% This function models exit of worst-performing agents that are replaced by inexperienced newcomers

function [theta1,theta2,adjust1,adjust2,perform,m1,h1] = ...
... noise_traders(perform,agents,features1,features2,theta1,theta2,adjust1,adjust2,m1,h1)

% Rate investors according to their wealth level
rating = [perform (1:agents)'];
lose = sortrows(rating,1);

% Prespecified number ofInvestors exit the market if their wealth falls below arbitrary threshold
avg_perform = sum(perform) / agents;
bankrupt = 3;
```

```
% New inexperienced agents take worst-performing agents' place
if ~isempty(perform < 0.7 * avg_perform)
   for i=1:bankrupt
      agent_index_lose1 = (((lose(i,2)-1) * size(features1,2)+ 1):((lose(i,2)-1) * size(features1,2) +
size(features1,2)))';
      agent_index_lose2 = (((lose(i,2)-1) * size(features2,2)+ 1):((lose(i,2)-1) * size(features2,2) +
size(features2,2)))';
      theta1(agent_index_lose1,:) = zeros(size(features1,2),size(theta1,2));
      theta2(agent_index_lose2,:) = zeros(size(features2,2),size(theta2,2));
      adjust1(lose(i,2)) = 1;
      adjust2(lose(i,2)) = 1;
   end
end
```

Appendix D. Program code for the parsimonious version of the model

Main program file

```
% Artificial stock market (ASM), populated by heterogeneous agents.
% Their behaviour is based on Q-learning with gradient-descent function
% approximation. Program implementation in MATLAB version R2007a.

clear all
load data_constant.mat                  % Load data describing controlled environment or real financial mkt.

% Model's key paremeters
stock_p       = 50;                      % Initial stock price
iterations    = length(next_div);        % Number of iterations
agents        = 200;                     % Number of agents
num_shares    = 100;                     % Number of shares
alpha         = 0.1;                     % Learning rate parameter
epsilon       = 0.1;                     % Exploration parameter
trade_cost    = 0;                       % Trade cost (as a fraction of transacted price)
for_horizon   = 10;                     % Earnings forecasting horizon, in quarters
loss_aversion = 1.3;                    % Loss aversion parameter, 1 is no loss aversion, >1 loss averse
penalty       = -2;                     % Penalty for the worst-performing agent


% Other model parameters
gamma1        = 0.1;                    % Exponential smoothing parameter
gamma2        = 0.1;                    % Exponential smoothing parameter
min_window    = 6;                      % Minimum historical sample size in earnings forecasting (quarters)
max_window    = 30;                     % Maximum historical sample size in earnings forecasting (quarters)
max_liquidity = 500;                    % Maximum amount of funds that individual agents can possess

% Technical initialisation
avg_d_p       = 0.01;                           % Exponentially weighted average of market price changes
d_p           = 0.01;                           % Change in market price
```

```
avg_p         = 50;                                % Exponentially weighted average of past market prices
exp_p         = ones(agents,1);                    % Vector of individual estimates of stock value
window        = min_window + ceil(rand(agents,1)...   % Vector of individual sample sizes
... * (max_window - min_window));                     % of earnings forecasting
orders        = [zeros(agents,2) (1:agents)'];     % Matrix containing individual trade orders
div           = ones(agents,1) * 0.015;            % Individual dividend payouts
earnings(:,1) = earnings(:,1) / num_shares;        % Vector of earnings per share (quarterly realisations)
dividend(:,1) = dividend(:,1) / num_shares;        % Vector of dividends per share (quarterly realisations)
exp_earnings  = earnings(1,1) * (1 + 0.1 * randn(500,agents)); % Matrix of corporate earnings forecasts
h             = [ones(num_shares,1);zeros(agents - num_shares,1)]; % Stock holdings before trading round
h1            = h;                                                  % Stock holdings after trading round
m             = [zeros(num_shares,1)+450;zeros(agents - num_shares,1) + 450 + stock_p];  % Cash holdings
                                                                    % before trading round
m1            = m;                                                  % Cash holdings after trading round
wealth        = h1 * stock_p + m1;                                  % Individual wealth at market prices
stock_p_prev  = stock_p;                                            % Last period's market stock price
returns       = ones(agents,1);                                     % Vector of individual returns on wealth
rewards       = zeros(agents,1);                                    % Vector of rewards
cum_rewards   = zeros(agents,1);                                    % Vector of cumulative rewards
fund          = ones(agents,1) * exp_earnings(1,1) * 4 / ytm(1,4);  % Risk-neutral estimates of fundamentals
w1            = zeros(agents,1) + 0;                                % Vector of RL adjustment factors
upper_bound   = 1000;                                               % Constraints on elements of RL theta matrix
%exp_earn     = ones(agents,1);

% Plotted variables and analytical indicators
plot_p          = zeros(iterations,1);
plot_div        = zeros(iterations,1);
plot_fund       = zeros(iterations,1);
plot_rewards    = zeros(iterations,2);
plot_avg_window = zeros(iterations,1);

% State features, experience matrix for Q-learning algorithm
num_features = 6;
features     = zeros(agents,num_features);
theta        = zeros(num_features * agents, 7);

% Starting actual agent behaviour; randomly decide on what action to take
action_index  = ceil(rand(agents,1) * 7);
action_matrix = [-0.02 -0.01 -0.005 0 0.005 0.01 0.02];
```

```
% Main cycle
for k = 1 : iterations

    % Update and store values of financial accounts (before trading round)
    h_prev = h;
    h = h1;
    m_prev = m;
    m = m1;

    % Determine individual forecasts of future corporate earnings
    index = find(earnings(:,2)==k, 1);
    if ~isempty(index)
        if index > 3
            window = selection (window,min_window,max_window,exp_earnings,earnings,cum_rewards,index,agents);
            [exp_earnings] = forecast (earnings, index, window, for_horizon, agents);
        else
            exp_earnings = earnings(index,1) * (1 + 0.1 * randn(500,agents));
        end
        cum_rewards = zeros(agents,1);
    else
        cum_rewards = cum_rewards + rewards;
    end
    fund_prev = fund;

    % Discount expected earnings to obtain individual risk-neutral fundamental valuations
    fund = zeros(agents,1);
    ttm = next_div(k) - 90;                         % Time to next earnings announcement (in days)
    for i = 1 : 500                                 % Discounting period is 500 quarters
        ttm = ttm + 90;
        if i == 1
            term = 1;
        elseif i >= length(quarters)
            term = length(quarters) - 1;
        else
            dif1 = abs(ttm - quarters(i));
            dif2 = abs(ttm - quarters(i+1));
            if dif1 < dif2
                term = i-1;
```

```
        else
            term = i;
        end
    end
    fund = (fund + exp_earnings(i,:)' * exp( -ytm(k,term) * ttm / 360)); % Iteratively compute individual
                                                                         % risk-neutral valuations of discounted earnings
end
fund(fund<0) = 0.01;

% Stochastic realisation of dividends on regular time intervals
if next_div(k) == 0
    div = dividend(index,1);
    plot_div(k) = div;
else
    div = 0;
    if k>1, plot_div(k) = plot_div(k-1); end
end
if next_div(k) == 0
    exp_div = exp_earnings(1,:)' * 0.4336;
    fund = fund - exp_div;
end
plot_fund(k) = mean(fund);

% Describe statistics of market dynamics (average price, trend)
avg_p = (1 - gamma2) * avg_p + gamma2 * stock_p;
d_p = stock_p - stock_p_prev;
avg_d_p = (1 - gamma1) * avg_d_p + gamma1 * d_p;
trend_p = stock_p + avg_d_p;

% Agents take actions determined by Q-learning algorithm to obtain individual assessments of stock value
w1 =  action_matrix(action_index)';
exp_p = fund ./ fund_prev .* stock_p .* (1 + w1);
exp_p(exp_p < 0) = 0;

% Determine individual trade orders
for i = 1 : agents
    orders(i,1) = exp_p(i);
    if (exp_p(i) > stock_p) && (m1(i) >= exp_p(i) * (1 + trade_cost)) && (h(i) == 0)
        orders(i,2) = 1;
```

```
    elseif (exp_p(i) < stock_p) &&  (h(i) == 1)
        orders(i,2) = -1;
    else
        orders(i,2) = 0;
    end
end

% Implement trading procedure
[orders_e,stock_p,stock_p_prev,turnover] = trade(orders,stock_p,stock_p_prev,agents,h,m,trade_cost);

% Update investors' accounts (after trading round)
h1 = h + orders_e(:,2);
m1_prev = m1;                                                % Store post-trading money balances
m1 = (m - orders_e(:,1) .* orders_e(:,2) - orders_e(:,1) .* abs(orders_e(:,2)) * trade_cost) ...
... * (exp( ytm(k,1) * 1 / 360)) + h1 * div;
wealth_prev = wealth;                                        % Store information on current wealth
wealth = h1 * stock_p + m1;

% Calculate returns on wealth and determine worst-performing agent
returns = wealth ./ wealth_prev - 1;
returns(returns<0) = returns(returns<0) * loss_aversion;
rating = [returns (1:agents)'];
lead = sortrows(rating,-1);
superior = zeros(agents,1);

% Calculate Q-learning rewards, penalize worst-performing agent
rewards = zeros(agents,1);
penalize = agents - find(lead(:,1)>lead(end,1),1,'last');
if isempty(penalize), penalize = agents; end
penalize = ceil(rand * penalize);
lead(end-penalize+1,1) = penalty;
rewards = sortrows(lead,2);
rewards(:,2) = [];

explore = zeros(agents,1);

% Observe resulting state of the world (state features)
f_prev = features;                                  % Store previous values of feature matrix
features(:,1) = ytm(k,4);
```

```
features(:,2) = fund /100;
features(:,3) = stock_p / 100;
features(:,4) = trend_p / stock_p;
features(:,5) = avg_p / stock_p;
features(:,6) = next_div(k) / 100;

% Implement learning procedure for each individual investor and decide on next action
for i = 1 : agents
    gamma = 0.995;
    agent_index = (((i-1) * num_features + 1) : (i * num_features))';   % Action reindexing
    Q_s_a = f_prev(i,:) * theta(agent_index,action_index(i));           % Estimate of Q-function
    delta = rewards(i) - Q_s_a;
    Q_s1_all = features(i,:) * theta(agent_index,:);      % Estimates of Q-functions of resultant state,
                                                          % associated with all possible actions
    max_Q_ind = find(Q_s1_all == max(Q_s1_all));          % Find action indices of largest Q-functions
    x = randperm(length(max_Q_ind));
    max_ind = max_Q_ind(x(1));
    max_Q = Q_s1_all(max_ind);
    delta = delta + gamma * max_Q;
    theta(agent_index,action_index(i)) = theta(agent_index,action_index(i)) + alpha* delta* f_prev(i,:)';
    if max(theta(agent_index,action_index(i))) > upper_bound
        theta(agent_index,:) = 0;
    end

    % Decide on next epsilon-greedy action
    if (rand > epsilon) && (k > 10)
        action_index(i) = max_ind;
    else
        action_index(i) = ceil(rand * 7);
        explore(i) = 1;
    end
end
returns = ones(agents,1);


m1(m1 > max_liquidity) = max_liquidity;                   % Remove excess liquidity
if next_div == 0, next_div = periods; end

% Update variables used for analytical purposes, display informative program execution indicators
```

```
    plot_p(k) = stock_p;
    plot_rewards(k,:) = [ rewards(1), mean(rewards) ];
    plot_weight(k,1:agents) = w1(1:agents);
    plot_avg_window(k) = mean(window);
    plot_turnover(k) = turnover;
    clc;
    disp([k, stock_p, plot_fund(k), plot_div(k), turnover, mean(w1)*100, plot_avg_window(k) ])
    disp('     k       stock_p     fund     div     turnover     w1      avg_window')

end

% Implement procedure of storing data on file (not reported here)
<…>
```

Function determining relevant historical data for earnings forecasts

```
% This function determines individual sample sizes of relevant historical earnings. This is done on the basis
% of random matching of agents, and, with prespecified probability, imitating the more successful agent's
% beliefs

function window = selection (window,min_window,max_window,exp_earnings,earnings,cum_rewards,index,agents)

epsilon = 0.1;
strategy = [ cum_rewards abs(exp_earnings(1,:) - earnings(index,1))' window (1:agents)' rand(agents,1) ];
match = sortrows(strategy,5);
for i = 1 : agents
    if rand > epsilon
        if strategy(i,1) < match(i,1)
            strategy(i,3) = match(i,3);
        end
    else
        strategy(i,3) = ceil(min_window + rand * (max_window - min_window));
    end
end
window = strategy(:,3);
```

Function describing formation of individual forecasts of corporate earnings

```
% This function computes matrix containing individual forecasts of corporate earnings. Computation is based
% on econometric models of trend reversion

function [ exp_earnings ] = forecast (series, index, window, for_horizon, agents)

exp_earnings = zeros(for_horizon,agents);

% Trend reversion, individual econometric estimation models for corporate earnings (levels)
y = series(1:index,1);
x = [ones(index,1) (1:index)'];
y_d1 = y(1:index) - [44.2; y(1:index-1)];
y_d2 = [44.2; y(1:index-1)] - [43.3; 44.2; y(1:index-2)];
mean_y_i = zeros(1,agents);

for i = 1 : agents
    % Individual error-correction models for differenced earnings
    if window(i) < index
        y_i = y(index-window(i)+1:index,1);
        x_i = [ones(window(i),1) (1:window(i))'];
        beta_i = x_i \ y_i;
        y_hat_i = x_i * beta_i;
        res_i = y_i - y_hat_i;

        y_d1_i = y_d1(index-window(i)+1:index);
        y_d2_i = y_d2(index-window(i)+1:index);
        y_sr_i = y_d1_i(2:window(i));
        x_sr_i = [y_d2_i(2:window(i)) res_i(1:window(i)-1)];
    else
        y_i = y;
        x_i = x;
        y_d1_i = y_d1;
        y_d2_i = y_d2;
        y_sr_i = y_d1_i(2:index);
        beta_i = x_i \ y_i;
        y_hat_i = x * beta_i;
        res_i = y - y_hat_i;
        x_sr_i = [y_d2_i(2:index) res_i(1:index-1)];
```

```
    end
    beta_sr_i = x_sr_i \ y_sr_i;
    y_hat_sr_i = x_sr_i * beta_sr_i;

    % Earnings forecasts based on the above regressions
    y_f_i = zeros (for_horizon,1);
    x_sr_f_i = zeros (for_horizon,2);

    y_f_i(1) = y_i(end) + y_hat_sr_i(end);
    x_sr_f_i(1,:) = [ y_f_i(1) - y_i(end), y_f_i(1) - (beta_i(1) + beta_i(2) * (index + 1))];
    y_f_i(2) = y_f_i(1) + x_sr_f_i(1,:) * beta_sr_i;
    x_sr_f_i(2,:) = [ y_f_i(2) - y_f_i(1), y_f_i(2) - (beta_i(1) + beta_i(2) * (index + 2))];


    for j = 3 : for_horizon
        y_f_i(j) = y_f_i(j-1) + x_sr_f_i(j-1,:) * beta_sr_i;
        x_sr_f_i(j,:) = [ y_f_i(j) - y_f_i(j-1), y_f_i(j) - (beta_i(1) + beta_i(2) * (index + i))];
    end
exp_earnings(:,i) = y_f_i;
mean_y_i(i) = mean(y_i);
end

% Form the matrix containing individual forecasts of future corporate earnings (500 quarters ahead)
mean_y_i(mean_y_i < 0) = 0;
exp_earnings = [exp_earnings; repmat(mean_y_i,500-for_horizon,1)];
exp_earnings(exp_earnings<0) = 0;
```

Function describing trading process

This function is conceptually identical to function *trade.m* described in Appendix C.

Tomas RAMANAUSKAS

STOCK MARKET SIMULATION BASED ON
INTERACTION OFHETEROGENEOUS LEARNING AGENTS

Doctoral Dissertation

Social Sciences,
Economics (04S)

Tomas RAMANAUSKAS

HETEROGENINIŲ BESIMOKANČIŲ AGENTŲ SĄVEIKA
PAGRĮSTAS IMITACINIS AKCIJŲ RINKOS MODELIAVIMAS

Daktaro disertacija

Socialiniai mokslai,
ekonomika (04S)

2010 04 28. 13,75 sp. l. Tiražas 20 egz.
Vilniaus Gedimino technikos universiteto
leidykla „Technika“,
Saulėtekio al. 11, 10223 Vilnius,
http://leidykla.vgtu.lt
Spausdino UAB „Biznio mašinų kompanija“,
J. Jasinskio g. 16A, 01112 Vilnius